

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:57:07 ; Search time 197 Seconds

(without alignments)

1413.611 Million cell updates/sec

Title: US-10-078-650-12

Perfect score: 484

Sequence: 1 MDEGIPHLQERLLEHRDFI.....PQNPSSAQEDPSQFGKEAP 494

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Uniprot 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	440	90.9	482	1 BHB3_HUMAN	Q9C0J9 homo sapien
2	339	70.0	482	2 O8TAT1	Q8TAT1 homo sapien
3	114	23.6	513	2 Q7YQC9	Q7YQC9 canis famil
4	81	16.7	193	2 Q8CGH4	Q8CGH4 mus musculu
5	81	16.7	410	1 BHB3_MOUSE	Q99PV5 mus musculu
6	81	16.7	410	1 BHB3_RAT	O35779 rattus norv
7	81	16.7	410	2 BADI8970	Badi8970 mus muscu
8	53	11.0	403	2 Q6NY50	Q6NY50 brachydanio
9	53	11.0	403	2 AAH66738	AAH66738 brachydan
10	53	11.0	409	2 Q6GNE9	Q6GNE9 xenopus lae
11	52	10.7	411	1 BHB2_MOUSE	O35185 mus musculu
12	52	10.7	411	1 BHB2_RAT	O35780 rattus norv
13	52	10.7	411	2 Q76JQ4	Q76JQ4 rattus norv
14	52	10.7	411	2 BAD01588	BAD01588 rattus no
15	52	10.7	411	2 BAC33281	BAC33281 mus muscu
16	52	10.7	411	2 BAC41056	BAC41056 mus muscu
17	52	10.7	412	1 BHB3_HUMAN	O14503 homo sapien
18	52	10.7	412	2 CAG33202	CAG33202 homo sapi
19	16	3.3	110	2 Q91WW0	Q91WW0 mus musculu
20	16	3.3	130	2 Q9TUC9	Q9TUC9 canis famil
21	16	3.3	130	2 Q9TST8	Q9TST8 felis silve
22	16	3.3	153	2 Q765Q0	Q765Q0 rattus norv
23	16	3.3	153	2 BAD08305	BAD08305 rattus no
24	16	3.3	166	2 Q7VPF4	Q7VPF4 anopheles g
25	16	3.3	167	2 Q765P9	Q765P9 rattus norv
26	16	3.3	167	2 BAD08306	BAD08306 rattus no
27	16	3.3	213	2 Q9TUC8	Q9TUC8 monodelphis
28	16	3.3	286	2 Q6F5E0	Q6F5E0 mus musculu
29	16	3.3	304	1 GSH2_HUMAN	Q9BZM3 homo sapien
30	16	3.3	305	1 GSH2_MOUSE	P31316 mus musculu
31	16	3.3	314	1 PMXB_MOUSE	Q99453 homo sapien

32	16	3.3	314	1 PMXB_MOUSE	O35690 mus musculu
33	16	3.3	314	2 Q6PUD9	Q6PUD9 homo sapien
34	16	3.3	314	2 AAH17199	AAH17199 homo sapi
35	16	3.3	315	2 Q6YW21	Q6YW21 oryza sativ
36	16	3.3	315	2 BAD05688	BAD05688 oryza sat
37	16	3.3	315	2 BAD05845	BAD05845 oryza sat
38	16	3.3	311	1 SIA2_BRARE	Q7SVL3 brachydanio
39	16	3.3	331	1 NK24_HUMAN	Q9H24 homo sapien
40	16	3.3	334	1 NK24_MOUSE	Q9EQM3 mus musculu
41	16	3.3	362	1 NK23_HUMAN	P97334 mus musculu
42	16	3.3	364	1 NK23_MOUSE	Q8TAU0 homo sapien
43	16	3.3	372	1 FBSH_HUMAN	Q9HAH7 homo sapien
44	16	3.3	376	1 FBSE_HUMAN	Q00358 homo sapien
45	16	3.3	378	1 FBSE_MOUSE	Q8R089 mus musculu

ALIGNMENTS

RESULT 1

BHB3_HUMAN
ID BHB3_HUMAN STANDARD; PRT; 482 AA.
AC Q9C0J9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (hDEC2) (Enhancer-of-split and
DE hairy-related protein 1) (SHARP-1).
GN Name=BHLHB3; Synonyms=DEC2, SHARP-1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=1162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
RN [2]
RP FUNCTION.
RX MEDLINE=21226716; PubMed=11278948; DOI=10.1074/jbc.M011619200;
RA Garriga-Canut M., Roopra A., Buckley N.J.;
RT "The basic helix-loop-helix protein, SHARP-1, represses transcription
RT by a histone deacetylase-dependent and histone deacetylase-independent
RT mechanism.";
RL J. Biol. Chem. 276:14821-14828(2001).
CC -!- FUNCTION: May be a transcriptional repressor that represses both
CC basal and activated transcription.
CC -!- SUBUNIT: Homodimerize.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and brain,
CC moderately expressed in pancreas and heart, weakly expressed in
CC placenta, lung, liver and kidney.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB044089; BAB21502.1; --
CC PIR; JC7583; JC7583.
CC Genbank; HGNC:16617; BHLHB3.
CC MIM; 606200; --
CC GO; GO:0005634; C:nucleus; NAS.

DR GO; GO:0003700; P:transcription factor activity; NAS.
 DR GO; GO:003054; P:cell differentiation; NAS.
 DR GO; GO:0008283; P:cell proliferation; NAS.
 DR GO; GO:0009887; P:organogenesis; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF07527; Hairly orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW DNA BIND 45
 KW DNA BINDING 57
 KW DNA BINDING 58
 KW DOMAIN 129
 KW DOMAIN 297
 FT DOMAIN 431
 FT DOMAIN 482 AA; 50497 MW; 2BECDC2FDEB7CE14 CRC64;
 SQ SEQUENCE 482 AA; 50497 MW; 2BECDC2FDEB7CE14 CRC64;
 Query Match 90.9%; Score 440; DB 1; Length 482;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025968; AAH25968.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF07527; Hairly orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00888; HLH; 1.
 SQ SEQUENCE 482 AA; 50525 MW; 23EBD420C6B0CE03 CRC64;
 Query Match 70.0%; Score 339; DB 2; Length 482;
 Best Local Similarity 99.8%; Pred. No. 3.8e-262;
 Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 104
 DB 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 102
 QY 105 TEQHQKIIALQNGERSLSKPSIQSDLDAPHSGFQTCAKEVLYLSRFESWTREPRCVQL 164
 DB 103 TEQHQKIIALQNGERSLSKPSIQSDLDAPHSGFQTCAKEVLYLSRFESWTREPRCVQL 162
 QY 165 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPVVIQ 224
 DB 163 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPVVIQ 222
 QY 225 RTQPSAELAAENDTDTDSYGGEAEARPDRKKGAGASRVTIKQPPGEDSPAPKMKL 284
 DB 223 RTQPSAELAAENDTDTDSYGGEAEARPDRKKGAGASRVTIKQPPGEDSPAPKMKL 282
 QY 285 DSRGGSGGGGGGAAAAAALLGPPAAAAALLRPDAALLSSLVAFGGGGGAPFPQAA 344
 DB 283 DSRGGSGGGGGGAAVAAAAALLGPPAAAAALLRPDAALLSSLVAFGGGGGAPFPQAA 342
 QY 345 AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKLYPAAAAAAPPFLLYPGIPAAAAAAA 404
 DB 343 AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKLYPAAAAAAPPFLLYPGIPAAAAAAA 402
 QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLIGAPHQHPGHRTHLPAGPRE 464
 DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLIGAPHQHPGHRTHLPAGPRE 462
 QY 465 PGNPSSAQEDPSQPGKEAP 484
 DB 463 PGNPSSAQEDPSQPGKEAP 482
 RESULT 2
 Q8TAT1 PRELIMINARY; PRT; 482 AA.
 ID Q8TAT1
 AC Q8TAT1
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Basic helix-loop-helix domain containing, class B, 3.
 GN Name=BHLHB3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa;
 OC Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettmann M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025968; AAH25968.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF07527; Hairly orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00888; HLH; 1.
 SQ SEQUENCE 482 AA; 50525 MW; 23EBD420C6B0CE03 CRC64;
 Query Match 70.0%; Score 339; DB 2; Length 482;
 Best Local Similarity 99.8%; Pred. No. 3.8e-262;
 Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 104
 DB 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 102
 QY 105 TEQHQKIIALQNGERSLSKPSIQSDLDAPHSGFQTCAKEVLYLSRFESWTREPRCVQL 164
 DB 103 TEQHQKIIALQNGERSLSKPSIQSDLDAPHSGFQTCAKEVLYLSRFESWTREPRCVQL 162
 QY 165 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPVVIQ 224
 DB 163 INHLHAVATQFLTPQLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPVVIQ 222
 QY 225 RTQPSAELAAENDTDTDSYGGEAEARPDRKKGAGASRVTIKQPPGEDSPAPKMKL 284
 DB 223 RTQPSAELAAENDTDTDSYGGEAEARPDRKKGAGASRVTIKQPPGEDSPAPKMKL 282
 QY 285 DSRGGSGGGGGGAAAAAALLGPPAAAAALLRPDAALLSSLVAFGGGGGAPFPQAA 344
 DB 283 DSRGGSGGGGGGAAVAAAAALLGPPAAAAALLRPDAALLSSLVAFGGGGGAPFPQAA 342
 QY 345 AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKLYPAAAAAAPPFLLYPGIPAAAAAAA 404
 DB 343 AAAPFCLPFCFLSPSAAAAAYVQFFLDKSGLEKLYPAAAAAAPPFLLYPGIPAAAAAAA 402
 QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLIGAPHQHPGHRTHLPAGPRE 464
 DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLIGAPHQHPGHRTHLPAGPRE 462
 QY 465 PGNPSSAQEDPSQPGKEAP 484
 DB 463 PGNPSSAQEDPSQPGKEAP 482

```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037742; AAH37742.1; -.
DR MGB; NGI1930704; Bhlhb3.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
DR NON_TER 193
FT SEQUENCE 193 AA; 21929 MW; 854890594639A422 CRC64;
SQ

Query Match 16.7%; Score 81; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLKSP 125
Db 103 TEQHQKIIALONGERSLKSP 123

RESULT 5
BHB3_MOUSE
ID BHB3_MOUSE STANDARD; PRT; 410 AA.
AC Q99PV5;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (mDEC2).
GN Name=Bhlhb3; Synonyms=dec2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092582; PubMed=11162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
CC -!- SUBUNIT: Homodimerize.
CC -!- FUNCTION: May be a transcriptional repressor that represses both
CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, brain and lung.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- SIMILARITY: Contains 1 orange domain.

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037742; AAH37742.1; -.
DR MGB; NGI1930704; Bhlhb3.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS50888; HLH; 1.
DR NON_TER 193
FT SEQUENCE 193 AA; 21929 MW; 854890594639A422 CRC64;
SQ

Query Match 23.6%; Score 114; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.1e-82;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTAL 102

QY 105 TEQHQKIIALONGERSLKSPISDLDAPHSQFQTCAKEVQLYLSRFESWTRE 158
Db 103 TEQHQKIIALONGERSLKSPISDLDAPHSQFQTCAKEVQLYLSRFESWTRE 156

RESULT 4
Q8CGH4
ID Q8CGH4 PRELIMINARY; PRT; 193 AA.
AC Q8CGH4;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Bhlhb3 protein (Fragment).
GN Name=Bhlhb3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Tosiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB044090; BAB21503.1; -
 DR PIR; JCV584; JCV584.
 DR MGD; MGI:1930704; Bhlhb3.
 DR InterPro; IPR001092; HLH basic.
 DR Pfam; PF07527; Hairy orange; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
 FT DNA BIND 45 57 Basic motif.
 FT DOMAIN 58 100 Helix-loop-helix motif.
 FT DOMAIN 129 175 Helix-loop-helix motif.
 FT DOMAIN 321 373 Helix-loop-helix motif.
 FT DOMAIN 321 373 Ala/Gly-rich.
 SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;

Query Match 16.7%; Score 81; DB 1; Length 410;
 Best Local Similarity 100.0%; Pred. No. 4e-56;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 DTYKLPRLIEKKRRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 104
 DB 43 DTYKLPRLIEKKRRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 102
 QY 105 TEQHQKIILALQNGERSLKSP 125
 DB 103 TEQHQKIILALQNGERSLKSP 123

RESULT 6
 ID_BHB3 RAT STANDARD; PRT; 410 AA.
 AC Q35779;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Enhancer-of-split
 DE and hairy-related protein 1) (SHARP-1).
 GN Names-Bhlhb3; Synonyms=Sharp;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Cerebellum;
 RX MEDLINE=98193761; PubMed=932582;
 RA Rossner M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;
 RT coupled to neuronal stimulation.
 RL Mol. Cell. Neurosci. 10:460-475(1997).
 CC -!- FUNCTION: May be a transcriptional repressor that represses both
 CC basal and activated transcription.
 CC -!- SUBUNIT: Homodimerize.
 CC -!- TISSUE SPECIFICITY: Highly expressed in subregions of the brain,
 CC moderately expressed in skeletal muscle, heart. Weakly expressed
 CC in lung.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- SIMILARITY: Contains 1 orange domain.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 249.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF009329; AAB63586.1; ALT_FRAME.
 DR RGD; 70900; Bhlhb3.
 DR InterPro; IPR001092; HLH basic.
 DR Pfam; PF07527; Hairy orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00511; ORANGE; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
 FT DNA BIND 45 57 Basic motif.
 FT DOMAIN 58 100 Helix-loop-helix motif.
 FT DOMAIN 129 175 Helix-loop-helix motif.
 FT DOMAIN 321 344 Ala/Gly-rich.
 SQ SEQUENCE 410 AA; 43917 MW; 829705CA3A013127 CRC64;

Query Match 16.7%; Score 81; DB 1; Length 410;
 Best Local Similarity 100.0%; Pred. No. 4e-56;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 DTYKLPRLIEKKRRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 104
 DB 43 DTYKLPRLIEKKRRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 102
 QY 105 TEQHQKIILALQNGERSLKSP 125
 DB 103 TEQHQKIILALQNGERSLKSP 123

RESULT 7
 ID_BAD18970 PRELIMINARY; PRT; 410 AA.
 AC BAD18970;
 DT 20-MAY-2004 (TrEMBLrel. 27, Created)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 20-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE BHLH transcriptional factor Dec2.
 GN BHLHB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=11162494;
 RX Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
 RA Yoshida E., Suardita K., Matsuda Y., Kato Y.;
 RT "Molecular cloning and characterization of DEC2, a new member of basic
 RT helix-loop-helix proteins."
 RL Biochem. Biophys. Res. Commun. 280:164-171(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=12397359;
 RA Honma S., Kawamoto T., Takagi Y., Fujimoto K., Sato F., Noshiro M.,
 RA Kato Y., Honma K.;
 RT "Dec1 and Dec2 are regulators of the mammalian molecular clock."
 RL Nature 419:841-844(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Hamaguchi H., Fujimoto K., Kawamoto T., Noshiro M., Maemura K.,
 RA Takeda N., Nagai K., Furukawa M., Honma S., Honma K., Kurihara H.,
 RA Kato Y.;
 RT "Expression of Dec2 -a basic helix-loop-helix transcription factor-
 RT gene is regulated by a molecular clock system."
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB126167; BAD18970.1; -.
SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;
Query Match 16.7%; Score 81; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 4e-56;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 104
DB 43 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 102

QY 105 TEQHQKIALLONGERSLKSP 125
DB 103 TEQHQKIALLONGERSLKSP 123

RESULT 8
Q6NY50
ID Q6NY50 PRELIMINARY; PRT; 403 AA.
AC Q6NY50;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:76913.
GN Name=zgc:76913;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.W.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RA Strausberg R.;
RA EMBL; BC066738; AAH66738.1; -.
RA Hypothetical protein.
DR EMBL; BC066738; AAH66738.1; -.
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FF298E CRC64;
Query Match 11.0%; Score 53; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.3e-34;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 97
DB 48 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 100

RESULT 9
AAH66738
ID AAH66738 PRELIMINARY; PRT; 403 AA.
AC AAH66738;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:76913.
GN ZGC:76913.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.W.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RA Strausberg R.;
RA EMBL; BC066738; AAH66738.1; -.
RA Hypothetical protein.
DR EMBL; BC066738; AAH66738.1; -.
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FF298E CRC64;
Query Match 11.0%; Score 53; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.3e-34;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 97
DB 48 DTYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTALKH 100

RESULT 10
Q6GNE9
ID Q6GNE9 PRELIMINARY; PRT; 409 AA.
AC Q6GNE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae: Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Stapleton L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Richards S., McEwan P.J., McKernan K.J., Casavant T.L., Scheetz T.E.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073563; AAH73563.1; -
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR003650; Orange.
 DR Pfam; PF07527; Hairy1_Orange; 1.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SMC0353; HLH; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 409 AA; 45446 MW; EECF92EC6845B0D7 CRC64;
 Query Match 11.0%; Score 53; DB 2; Length 409;
 Best Local Similarity 100.0%; Pred. No. 9.4e-34;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 DTYKLPRLIEKKRRDRINECTAQLKDLPEHLKLTTLGLHLEKAVVLETLKH 97
 Db 93 DTYKLPRLIEKKRRDRINECTAQLKDLPEHLKLTTLGLHLEKAVVLETLKH 145
 RESULT 11
 BHB2_MOUSE STANDARD; PRT; 411 AA.
 AC C35155; P97289;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Class B basic helix-loop-helix protein 2 (bHLHB2) (stimulated with
 DE retinoic acid 13) (E47 interaction protein 1) (eip1).
 GN Name=bHLhb2; Synonyms=Stral3, Clast5;
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic carcinoma;
 RX MEDLINE=97427857; PubMed=9284045;
 RA Boudjelal M., Taneja R., Matsubara S., Bouillet P., Dollé P.,
 RA Chambon P.;
 RT "Overexpression of Stral3, a novel retinoic acid-inducible gene of the
 RT basic helix-loop-helix family, inhibits mesodermal and promotes
 RT neuronal differentiation of P19 cells.";
 RL Genes Dev. 11:2052-2065(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH E47/TCF3.
 RC STRAIN=BALB/C;
 RX MEDLINE=97203270; PubMed=9050988; DOI=10.1038/sj.onc.1200912;
 RA Dear T.N., Hainzl T., Folio M., Nehls M., Wilmore H., Matena K.,
 RA Boehm T.;
 RT "Identification of interaction partners for the basic-helix-loop-helix
 RT protein E47.";
 RL Oncogene 14:891-898(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Spleen;
 RA O-Wang J.;
 RT "Isolation of a CD40-activated gene from murine splenic B cells.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Stapleton L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May function as a transcriptional factor for neuronal
 CC differentiation.
 CC -!- SUBUNIT: Able to homodimerize or heterodimerize with E47.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- DEVELOPMENTAL STAGE: Expressed from E9.5 day to E17.5 day in the
 CC ventricular layer of the brain and spinal cord, but also in the
 CC retinal pigment epithelium, developing eyelids, nasal epithelium,
 CC serous gland, vibrissae, epithelium of the mouth cavity and the
 CC tooth buds. Highly expressed in the heart, thymus and adrenal
 CC glands followed by lung, liver parenchyma, kidney tubules,
 CC epithelium of the esophagus and stomach. From E15.5 to E17.5 it is
 CC expressed in urinary bladder and urethra. From E17.5, it is
 CC expressed in developing muscle.
 CC -!- INDUCTION: Stimulated by retinoic acid (RA).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- SIMILARITY: Contains 1 orange domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch)

EMBL; AF009330; AAB63587.1; -.
RGD; 68439; Ehlnb2.
InterPro; IPR001092; HLH_basic.
InterPro; IPR003650; Orange.
Pfam; PF07527; Hairy orange; 1.
Pfam; PF00010; HLH; 1.
SMART; SM00353; HLH; 1.
SMART; SM00511; ORANGE; 1.
PROSITE; PS50888; HLH; 1.
DNA-binding; Nuclear protein; Transcription regulation.
DNA_BIND 53 65
DOMAIN 66 108
HELM 140 184
SEQUENCE 411 AA; 45529 MW; E56BD468D08924AD CRC64;
Orange.

```
Query Match      10.7%; Score 52; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 13	
Q75U04	
AC Q76J04	PRELIMINARY;
ID Q76J04	PRT; 411 AA.
CD Q76J04	
DT 05-JUL-2004	(TRENDEL, 27, Created)
DT 05-JUL-2004	(TRENDEL, 27, Last sequence update)
DT 05-JUL-2004	(TRENDEL, 27, Last annotation update)
DT 05-JUL-2004	(TRENDEL, 27, Last annotation update)
DE Enhancer of	split- and hairy-related protein-3

[illegible]

```

InterPro: IPR001092; HLH-basic.
InterPro: IPR003650; Orange.
Pfam: PF07527; Hairly_orange; 1.
Pfam: PF00010; HLH; 1.
SMART: SMO0353; HLH; 1.
SMART: SMO0511; Orange; 1.
PROSITE: PS00888; HLH; 1.
SEQUENCE 411 AA; 45557 MW; 2086DAB83DB82058 CRC64;

Query Match      10.7%; Score 52; DB 2; Length 411;
Best Local Similarity 100.0%; Pred.No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

[illegible]

```

DN 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split- and hairy-related protein-2.
GN SHARP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Hirano S., Yamada K., Kawata H., Shou Z., Mizutani T., Shigematsu Y.,
RA Mayumi M., Miyamoto K.;
RT "The rat enhancer of split- and hairy-related protein-2 gene: hepatic
RT expression, genomic structure, and promoter analysis.";
RL Arch. Biochem. Biophys. 422:81-90(2004).
DR EMBL: AB096137; BAD01588.1; -.
SQ SEQUENCE 411 AA; 43557 MW; 2086DABB3DB82058 CRC64;

Query Match 10.7%; Score 52; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTIKH 97
DB 52 TYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTIKH 103

RESULT 15
ID BAC33281 PRELIMINARY; PRT; 411 AA.
AC BAC33281;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE 16 days embryo head cDNA, RIKEN full-length enriched library,
DE clone: C130042M06 product: basic helix-loop-helix domain containing,
DE class B2, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=2049374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK048242; BAC33281.1; -.
SQ SEQUENCE 411 AA; 45360 MW; B392893CD49292BC CRC64;

Query Match 10.7%; Score 52; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTIKH 97
DB 52 TYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTIKH 103

```

Search completed: December 15, 2004, 20:14:18
Job time : 199 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:56:37 ; Search time 155 Seconds
(without alignments)
1120.161 Million cell updates/sec

Title: US-10-078-650-12

Perfect score: 484

Sequence: 1 MDEGPHLQRLLEHRDFI.....PGNPSSAQEDPSQPGKEAP 484

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1980s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	100.0	484	4 AAB70693	Aab70693 Human DEC
2	440	90.9	482	4 AAB70692	Aab70692 Human DEC
3	440	90.9	482	5 ABG96308	Abg96308 Human ova
4	440	90.9	482	6 ABO14645	Abol14645 Novel hum
5	440	90.9	482	6 ABR58470	Abr58470 Human BHL
6	440	90.9	482	7 ADC31646	Adc31646 Human nov
7	440	90.9	482	7 ADES4662	Ades4662 Human Pro
8	440	90.9	482	8 ADN05014	Adn05014 Antipsori
9	440	90.9	482	5 ABJ04644	Abj04644 Protein o
10	159	32.9	165	6 AAU16188	Aau16188 Human nov
11	159	32.9	165	6 ABUS5257	Abus5257 Human nov
12	108	22.3	203	6 ABO14646	Abol14646 Novel hum
13	107	22.1	150	4 AAU16603	Aau16603 Human nov
14	107	22.1	150	6 ABUS5672	Abus5672 Human nov
15	81	16.7	410	4 AAB70694	Aab70694 Mouse DEC
16	81	16.7	410	7 ADES4660	Ades4660 Rat Prote
17	76	15.7	88	4 AAU32415	Aau32415 Novel hum
18	62	12.8	109	5 ADK34819	Adk34819 Novel hum
19	54	11.2	54	5 ABP52802	Abp52802 Medane an
20	52	10.7	379	8 ADP46580	Adp46580 Human col
21	52	10.7	379	8 ADP46579	Adp46579 Human col
22	52	10.7	404	5 ABB57218	Abb57218 Mouse isc
23	52	10.7	411	7 ADES4664	Ades4664 Rat Prote
24	52	10.7	411	7 ADD90602	Add90602 Mouse BHL
25	52	10.7	411	8 ADP74022	Adp74022 Murine DE

26	52	10.7	412	2 AAW99060	Aaw99060 Human bas
27	52	10.7	412	5 ABP65203	Abp65203 Hypoxia-r
28	52	10.7	412	7 ADES4666	Ades4666 Human Pro
29	52	10.7	412	7 ADD90600	Add90600 Human BHL
30	52	10.7	412	8 ADL82901	Adl82901 Human PRO
31	52	10.7	412	8 ADN03704	Adn03704 Antipsori
32	52	10.7	412	8 ADO19143	Ado19143 Human PRO
33	52	10.7	412	8 ADO19145	Ado19145 Human PRO
34	52	10.7	425	8 ADP46583	Adp46583 Human col
35	52	10.7	425	8 ADP46584	Adp46584 Human col
36	49	10.1	207	4 AAU30778	Aau30778 Novel hum
37	47	9.7	54	5 ABP52803	Abp52803 Medane an
38	43	8.9	412	2 AAW08140	Aaw08140 Human cyt
39	43	8.9	412	3 AAY87958	Aay87958 Human CR8
40	43	8.9	412	4 AAB90758	Aab90758 Human she
41	42	8.7	204	7 ADC33242	Adc33242 Human nov
42	28	5.8	432	4 AAU30779	Aau30779 Novel hum
43	16	3.3	18	8 ADH58940	Adh58940 Silk prot
44	16	3.3	19	7 AAO30392	Aao30392 Human ARX
45	16	3.3	20	4 AAB69489	Aab69489 Synthetic

ALIGNMENTS

RESULT 1

AAB70693
ID AAB70693 standard; protein; 484 AA.

XX AC AAB70693;

XX DT 17-MAY-2001 (first entry)

XX DE Human DEC2b protein sequence SEQ ID NO:12.

XX KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;

XX KW basic helix loop helix protein; cell differentiation; proliferation.

XX OS Homo sapiens.

XX FN WO200114551-A1.

XX PD 01-MAR-2001.

XX PF 19-JUN-2000; 2000WO-JP003991.

XX PR 19-AUG-1999; 99JP-00233286.

XX (CHUS) CHUGAI SEIYAKU KK.

XX PI Fujimoto K, Shin M, Kato Y;

XX DR WPI; 2001-202935/20.

XX DR N-PSDB; AAF74776.

XX DEC2 is a basic helix loop helix protein of the DEC family for use in development of drugs for treatment of disorders of cell differentiation and proliferation.

XX PS Claim 2; Page 69-72; 83pp; Japanese.

XX The present invention describes a basic helix loop helix (bHLH) type transcription factor designated DEC2. DEC2 can be used as a tool in the development of drugs for the treatment and prevention of disorders involving cell differentiation and proliferation. The present sequence represents the specifically claimed human DEC2b protein, as given in the present invention

XX SQ Sequence 484 AA;

Query Match 100.0%; Score 484; DB 4; Length 484;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDEGIPHLQEROLLEHRRDFTGLDYSLYMCKPKRSMKRDRTKVSQTYKLPHELIEKKRRD 60
DB 1 MDEGIPHLQEROLLEHRRDFTGLDYSLYMCKPKRSMKRDRTKVSQTYKLPHELIEKKRRD 60
QY 61 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLETLTKHLKALTALTEQOQKIIALONGER 120
DB 61 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLETLTKHLKALTALTEQOQKIIALONGER 120
QY 121 SLKSPITQSDLDAPHSGFQTCACEVLYLGRFSWTPREPRCVQLINHLHAVATQELPTPQ 180
DB 121 SLKSPITQSDLDAPHSGFQTCACEVLYLGRFSWTPREPRCVQLINHLHAVATQELPTPQ 180
QY 181 LLTQOVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPIQRTQPSAEIAAENDTDT 240
DB 181 LLTQOVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPIQRTQPSAEIAAENDTDT 240
QY 241 DSCYGGGEAEARPDREKKGAGASRVTIKQPPGEDSPAPKMKLDSRGGSGGGGGGAA 300
DB 241 DSCYGGGEAEARPDREKKGAGASRVTIKQPPGEDSPAPKMKLDSRGGSGGGGGGAA 300
QY 301 AAAAAALLGPDPAALALLRDPDAALLSSLVAFGGGGGAPFPQAAAAAPCLPFCFLSPSA 360
DB 301 AAAAAALLGPDPAALALLRDPDAALLSSLVAFGGGGGAPFPQAAAAAPCLPFCFLSPSA 360
QY 361 AAAYVQFPLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAALAAAAAALPCLSSV 420
DB 361 AAAYVQFPLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAAALAAAAAALPCLSSV 420
QY 421 LSPPEKAGAAAATLLPHEVAPLGPAPHPQHPHGRTHLPFAGPREPKNPSSAQEDPSQPG 480
DB 421 LSPPEKAGAAAATLLPHEVAPLGPAPHPQHPHGRTHLPFAGPREPKNPSSAQEDPSQPG 480
QY 481 KEAP 484
DB 481 KEAP 484

RESULT 2
ID AAB70692 standard; protein; 482 AA.
AC AAB70692;
XX
XX
XX 17-MAY-2001 (first entry)
DE Human DEC2a protein sequence SEQ ID NO:2.
XX
XX DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
XX Homo sapiens.
XX
XX WO200114551-A1.
XX
XX 01-MAR-2001.
XX
XX 19-JUN-2000; 2000WO-JP003991.
XX
XX 19-AUG-1999; 99JP-00233286.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Fujimoto K, Shin M, Kato Y;
XX
XX WPI; 2001-202935/20.
XX
XX N-PSDB; AAF74767.
XX
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
XX development of drugs for treatment of disorders of cell differentiation
XX and proliferation.
XX
XX Claim 2; Page 55-59; 83pp; Japanese.
XX
PS
```

```
XX The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed human DEC2a protein, as given in the
CC present invention
XX
SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 4; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLTGHLKAVVLETLTKHLKALTAL 104
DB 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLTGHLKAVVLETLTKHLKALTAL 102
QY 105 TEQOQKIIALONGERSLKSPISQSDLDAPHSGFQTCACEVLYLSRFESWTPREPRCVQL 164
DB 103 TEQOQKIIALONGERSLKSPISQSDLDAPHSGFQTCACEVLYLSRFESWTPREPRCVQL 162
QY 165 INHLHAVATQELPTPQLLTQOVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPIQ 224
DB 163 INHLHAVATQELPTPQLLTQOVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCVPIQ 222
QY 225 RTQPSAEIAAENDTDTDSYGGGEAEARPDREKKGAGASRVTIKQPPGEDSPAPKMKL 284
DB 223 RTQPSAEIAAENDTDTDSYGGGEAEARPDREKKGAGASRVTIKQPPGEDSPAPKMKL 282
QY 285 DSRGGSGGGGGGAAAAAALLGPDPAALALLRDPDAALLSSLVAFGGGGGAPFPQAAA 344
DB 283 DSRGGSGGGGGGAAAAAALLGPDPAALALLRDPDAALLSSLVAFGGGGGAPFPQAAA 342
QY 345 AAAPFCLPFCFLSPSAAGSAAYVQFPLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAA 404
DB 343 AAAPFCLPFCFLSPSAAGSAAYVQFPLDKSGLEKLYLPAAAAAPFLLYPGIPAPAAAAA 402
QY 405 AAAAAAALPCLSSVLSPPPEKAGAAAATLLPHEVAPLGPAPHPQHPHGRTHLPFAGPRE 464
DB 403 AAAAAAALPCLSSVLSPPPEKAGAAAATLLPHEVAPLGPAPHPQHPHGRTHLPFAGPRE 462
QY 465 PGNPSSAQEDPSQPGKEAP 484
DB 463 PGNPSSAQEDPSQPGKEAP 482

RESULT 3
ABG96308
ID ABG96308 standard; protein; 482 AA.
XX
XX AC ABG96308;
XX
XX DT 11-DEC-2002 (first entry)
XX
XX DE Human ovarian cancer marker OV4.
XX
XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
XX OS Homo sapiens.
XX
XX FN WO200271928-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US007826.
XX
XX
```

PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324567P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Monahar JE, Gannavakapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR N-PSDB; ABS76400.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX Disclosure; Page 187-188; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. non-tuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention
XX
SQ Sequence 482 AA;

Query Match 90.9%; Score 440; DB 5; Length 482;
Best Local Similarity 100.0%; Pred. NC. 0;
Matches 440; Conservative 0; Mismatches 0; Gaps 0;

QY 45 DYTKLPHRIEKKRRRINECTAQLKDLPEHLKLTTLGLHKA VVLETLKHLKALNAL 104
DB 43 DYTKLPHRIEKKRRRINECTAQLKDLPEHLKLTTLGLHKA VVLETLKHLKALNAL 102

QY 105 TEQHQKIALQNGERSLKSPQISQDLDAFHSQFQTCACEVLYLGRFESWTREPRCVQL 164
DB 103 TEQHQKIALQNGERSLKSPQISQDLDAFHSQFQTCACEVLYLGRFESWTREPRCVQL 162

QY 165 INHLHAVAQFLPTPOLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCPVVIQ 224
DB 163 INHLHAVAQFLPTPOLLTQQVPLSKGTGAPSAAGSAAPCLERAGOKLEPLAYCPVVIQ 222

QY 225 RTOPSAELAENDTDTDSYGGGEAEARPDRKXGKAGASRVTIKQEPGDSPPAPKMKL 284
DB 223 RTOPSAELAENDTDTDSYGGGEAEARPDRKXGKAGASRVTIKQEPGDSPPAPKMKL 282

QY 285 DSRGGSGGGPGGGGAAAAAALLGPDPAFAAAALLRDPDAALLSSLVAFGGGGGAPFPQAAA 344
DB 283 DSRGGSGGGPGGGGAAAAAALLGPDPAFAAAALLRDPDAALLSSLVAFGGGGGAPFPQAAA 342
QY 345 AAAPFCFLPFCFLSPSAAAAVQPLDKSGLEKLYPAAAAAPPELLLYPGIPAPAAAAAAA 404
DB 343 AAAPFCFLPFCFLSPSAAAAVQPLDKSGLEKLYPAAAAAPPELLLYPGIPAPAAAAAAA 402
QY 405 AAAAAAFAFPCLSSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPHQHPGHRTHLPFAGPRE 464
DB 403 AAAAAAFAFPCLSSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPHQHPGHRTHLPFAGPRE 462
QY 465 PCNPESQAQEDPSQPGKEAP 484
DB 463 PCNPESQAQEDPSQPGKEAP 482

RESULT 4
ABO14645
ID ABO14645 standard; protein; 482 AA.
XX
AC ABO14645;
XX
DT 25-AUG-2003 (first entry)
XX
DE Novel human protein #18.
XX
KW Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection.
XX
OS Homo sapiens.
XX
FN WO2003023002-A2.
XX
PD 20-MAR-2003.
XX
PF 09-SEP-2002; 2002WO-US028539.
XX
PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 17-SEP-2001; 2001US-0322817P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324980P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DM, Zhong M;
PI Gerlach VL, Vernet CAM, Berghs C, Rotherberg ME, Guo X;
PI Shmukets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shency SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
XX WPI; 2003-313242/30.
DR N-PSDB; ACD19338.
XX

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
XX stroke or infections.
XX
PS Claim 1; Page 125; 586pp; English.
XX
XX The invention describes a new isolated polypeptide (NOVX). The NOVX
XX polypeptide, nucleic acid and antibody are useful as therapeutics,
XX particularly in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, which includes a pathology associated
XX with NOVX polypeptide. The DNA encoding the protein is useful in gene
XX therapy for treating the disease or condition. In particular, the NOVX
XX polypeptide or polynucleotide is useful for treating endocrine/
XX metabolism-related diseases (e.g. obesity or diabetes), central nervous
XX system disorders (e.g. Alzheimer's disease, Parkinson's disease,
XX epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
XX and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
XX asthma, inflammatory bowel disease, rheumatoid arthritis or
XX osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
XX prostate or brain cancers, or melanoma), liver diseases (e.g. liver
XX cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
XX haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
XX These are also useful in developing powerful assay system for functional
XX analysis of various human disorders, as well as in diagnostic
XX applications, and for monitoring the effects of drugs during clinical
XX trials. This is the amino acid sequence of a novel human NOV protein
XX
XX Sequence 482 AA;

Query Match 90.9%; Score 440; DB 6; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLELTTLKHLKALTAL 104
Db 43 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLELTTLKHLKALTAL 102
QY 105 TEQCHOKITLQNGERSLSKPSIQSDLDAPHSGFQTCAGEVQLYLSRFESWTPREPCVQL 164
Db 103 TEQCHOKITLQNGERSLSKPSIQSDLDAPHSGFQTCAGEVQLYLSRFESWTPREPCVQL 162
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
Db 163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAEADTDDTDSGYGGEAEAPDREKKGAGASRVTIKQPPGEDSPAPKMKL 284
Db 223 RTQPSAELAEADTDDTDSGYGGEAEAPDREKKGAGASRVTIKQPPGEDSPAPKMKL 282
QY 285 DSRGGSGGGPGGGAAGAAALLLGPDPAALLLRDPDAALLSLVAFGGGGGAPFPQAA 344
Db 283 DSRGGSGGGPGGGAAGAAALLLGPDPAALLLRDPDAALLSLVAFGGGGGAPFPQAA 342
QY 345 AAPFCPLPFCFLSFAAAAAYVQFLDKSGLEKLYLPAAAAAPFLYPCIPAPAAAAAAA 404
Db 343 AAPFCPLPFCFLSFAAAAAYVQFLDKSGLEKLYLPAAAAAPFLYPCIPAPAAAAAAA 402
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLGAPHPQHPGHTLPPFAGPRE 464
Db 403 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLGAPHPQHPGHTLPPFAGPRE 462
QY 465 PGNPSSAQEDPSQPGKEAP 484
Db 463 PGNPSSAQEDPSQPGKEAP 482

RESULT 5
ABR58470
ID ABR58470 standard; protein; 482 AA.
XX
AC ABR58470;
XX

DT 07-JUL-2003 (first entry)
XX
DE Human bHLH protein DEC2 from ovarian cancer clone O1668P.
XX
KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.
XX
OS Homo sapiens.
XX
FN WO2003029468-A1.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031467.
XX
PR 02-OCT-2001; 2001US-0327135P.
XX
PR 30-MAY-2002; 2002US-0384531P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2003-372001/35.
XX
PT New polynucleotide and polypeptide useful for diagnosing and/or treating
XX cancer, particularly ovarian cancer, and as a vaccine.
XX
PS Claim 2; Page 156-157; 169pp; English.
XX

CC The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention
XX
XX Sequence 482 AA;

Query Match 90.9%; Score 440; DB 6; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLELTTLKHLKALTAL 104
Db 43 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGHLKAVVLELTTLKHLKALTAL 102
QY 105 TEQCHOKITLQNGERSLSKPSIQSDLDAPHSGFQTCAGEVQLYLSRFESWTPREPCVQL 164
Db 103 TEQCHOKITLQNGERSLSKPSIQSDLDAPHSGFQTCAGEVQLYLSRFESWTPREPCVQL 162
QY 165 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 224
Db 163 INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAEADTDDTDSGYGGEAEAPDREKKGAGASRVTIKQPPGEDSPAPKMKL 284
Db 223 RTQPSAELAEADTDDTDSGYGGEAEAPDREKKGAGASRVTIKQPPGEDSPAPKMKL 282
QY 285 DSRGGSGGGPGGGAAGAAALLLGPDPAALLLRDPDAALLSLVAFGGGGGAPFPQAA 344
Db 283 DSRGGSGGGPGGGAAGAAALLLGPDPAALLLRDPDAALLSLVAFGGGGGAPFPQAA 342
QY 345 AAPFCPLPFCFLSFAAAAAYVQFLDKSGLEKLYLPAAAAAPFLYPCIPAPAAAAAAA 404
Db 343 AAPFCPLPFCFLSFAAAAAYVQFLDKSGLEKLYLPAAAAAPFLYPCIPAPAAAAAAA 402
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLGAPHPQHPGHTLPPFAGPRE 464
Db 403 AAAAAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLGAPHPQHPGHTLPPFAGPRE 462
QY 465 PGNPSSAQEDPSQPGKEAP 484
Db 463 PGNPSSAQEDPSQPGKEAP 482

RESULT 6
AD31646
ID AD31646 standard; protein; 482 AA.
XX AC AD31646;
XX AC AD31646;
XX 18-DEC-2003 (first entry)
XX Human novel polypeptide sequence, SEQ ID NO:1728.
XX Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX ulcers; osteoporosis; autoimmune disease; cancer;
XX molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
XX anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX gene therapy.
XX Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSB-) HYSB INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI: 2003-371981/35.
XX N-PSDB; ADC30675.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
XX treating conditions such as neurodegenerative diseases, anemias, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
XX Claim 20; SEQ ID NO 1728; 1185pp; English.
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX contig sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,

CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 7; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTUKHLKALTAL 104
DB 43 DTYKLPRLIEKKRRDRINECIAQKDLLPEHLKLTTLGHLEKAVVLELTUKHLKALTAL 102
QY 105 TEQHQKIITIALQNGERSLKSPISQSDLDAPHSFGFOTCAKEVLYLSRFSMTPEPRCVOL 164
DB 103 TEQHQKIITIALQNGERSLKSPISQSDLDAPHSFGFOTCAKEVLYLSRFSMTPEPRCVOL 162
QY 165 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 222
QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGEDSPAPKMKL 284
DB 223 RTQPSAELAAENDTDTDSGYGGEAEARPDREKKGAGASRVTIKQEPPEGEDSPAPKMKL 282
QY 285 DSRGGSGGGGGGAAAAAALLGPDPAALALRDPDAALLSSLVAFGGGGGAPFPQAA 344
DB 283 DSRGGSGGGGGGAAAAAALLGPDPAALALRDPDAALLSSLVAFGGGGGAPFPQAA 342
QY 345 AAAPFCFLPFCFLSPSAAAAVVOPELDSKGLSKLYPAAAAAPFLLYPGIPAPAAAAAAA 404
DB 343 AAAPFCFLPFCFLSPSAAAAVVOPELDSKGLSKLYPAAAAAPFLLYPGIPAPAAAAAAA 402
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPHQPQHPGHRTHLPFAGPRE 464
DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPHQPQHPGHRTHLPFAGPRE 462
QY 465 PGNPSSAQEDPSPGKKEAP 484
DB 463 PGNPSSAQEDPSPGKKEAP 482
RESULT 7
ADE54662
ID ADE54662 standard; protein; 482 AA.
XX ADE54662;
XX
XX 29-JAN-2004 (first entry)
XX Human Protein Q9C0J9, SEQ ID NO 467.
XX
XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-033347P.
XX (GEBO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX

XX PI Woelf C, D'urso D, Befort K, Costigan M;
XX DR WPI; 2003-268312/26.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 7; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLLEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 104
DB 43 DTYKLPRLLEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 102
QY 105 TEQHQKIIALQNGERSLKSPISQDLDAPHSGFQTCAKEVLQYLSRFSWTPREPCVOL 164
DB 103 TEQHQKIIALQNGERSLKSPISQDLDAPHSGFQTCAKEVLQYLSRFSWTPREPCVOL 162
QY 165 INHLHAVATQFLPTPQLLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPQLLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 222
QY 225 RTQPSAEALAEANDTDTDSYGGEAEARPDREKKGAGASRVTIKQEPGDSPPAKEMKL 284
DB 223 RTQPSAEALAEANDTDTDSYGGEAEARPDREKKGAGASRVTIKQEPGDSPPAKEMKL 282
QY 285 DSRGGSGGGGGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 344
DB 283 DSRGGSGGGGGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 342
QY 345 AAPFCLPFCFLSPSAAAAVQPFLLDKSGLEKLYLFAAAAAAPPELLYPGIPAPAAAAA 404
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPQHPHGRTHLPFAGPRE 464
DB 403 AAAAAAAPPCLSSVLSPPPEKAGAAAAATLLPHEVAPLGAHPQHPHGRTHLPFAGPRE 462
QY 465 PGNPESQAQEDPSQPGKEAP 484

DB 463 PGNPESQAQEDPSQPGKEAP 482
RESULT 8
ADN05014
ID ADN05014 standard; protein; 482 AA.
XX AC ADN05014;
XX DI 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #687.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX FN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX PI Wu TD;
XX DR WPI; 2004-305105/28.
XX DR N-PSDB; ADN05013.
XX PT New PRO nucleic acid or polypeptide, useful for preparing a
XX PT pharmaceutical composition for diagnosing or treating psoriasis in a
XX PT mammal.
XX PS Claim 9; SEQ ID NO 1408; 3069pp; English.
XX CC The invention relates to novel polynucleotide and polypeptides for
XX CC treating psoriasis or a sequence having at least 80% identity to the
XX CC above sequences. The nucleic acid is useful for preparing a composition
XX CC for diagnosing or treating psoriasis in a mammal. This sequence
XX CC corresponds to one of the polypeptides of the invention.
XX SQ Sequence 482 AA;
Query Match 90.9%; Score 440; DB 8; Length 482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLLEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 104
DB 43 DTYKLPRLLEKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHKALTAL 102
QY 105 TEQHQKIIALQNGERSLKSPISQDLDAPHSGFQTCAKEVLQYLSRFSWTPREPCVOL 164
DB 103 TEQHQKIIALQNGERSLKSPISQDLDAPHSGFQTCAKEVLQYLSRFSWTPREPCVOL 162
QY 165 INHLHAVATQFLPTPQLLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 224
DB 163 INHLHAVATQFLPTPQLLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 222
QY 225 RTQPSAEALAEANDTDTDSYGGEAEARPDREKKGAGASRVTIKQEPGDSPPAKEMKL 284
DB 223 RTQPSAEALAEANDTDTDSYGGEAEARPDREKKGAGASRVTIKQEPGDSPPAKEMKL 282
QY 285 DSRGGSGGGGGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 344
DB 283 DSRGGSGGGGGGAAAAAALLGPDPAALLLRPDAALLSSLVAFGGGGGAPFPQPA 342
QY 345 AAPFCLPFCFLSPSAAAAVQPFLLDKSGLEKLYLFAAAAAAPPELLYPGIPAPAAAAA 404

Db 343 AAAPFCFLPFCFLSPSAAAAYVQFLDKSGLEKLYFAAAAAPFLLYPGIPAPAAAAAA 402
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAATLTPHEVAPLGAHPHQPHQTHLPAGPRE 464
Db 403 AAAAAAAPPCLSSVLSPPPEKAGAAAATLTPHEVAPLGAHPHQPHQTHLPAGPRE 462
QY 465 PGNPSSAQEDPSQPGKEAP 484
Db 463 PGNPSSAQEDPSQPGKEAP 482

RESULT 9
ABJ04644
ID ABJ04644 standard; protein; 767 AA.
AC ABJ04644;
XX
DT 11-OCT-2002 (first entry)
XX
DE Protein of NOVX 5 SEQ ID No 10.
XX

KW Cytostatic; antidiabetic; anorectic; metabolic; neurotropic; antilipemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenia; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorders;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW gene therapy.
XX

OS Unidentified.
XX

XX WO200246409-A2.
PN

XX 13-JUN-2002.
XX

XX 06-DEC-2001; 2001WO-US046586.
XX

PR 06-DEC-2000; 2000US-0251660P.
PR

PR 12-DEC-2000; 2000US-0255029P.
PR

PR 08-JAN-2001; 2001US-0260326P.
PR

PR 24-JAN-2001; 2001US-0263800P.
PR

PR 20-FEB-2001; 2001US-0269942P.
PR

PR 24-APR-2001; 2001US-0286183P.
PR

PR 20-AUG-2001; 2001US-0313627P.
PR

PR 12-SEP-2001; 2001US-0318712P.
PR

XX (CURA-) CURAGEN CORP.
XX

XX Guo X, Li L, Patturajan M, Shinkets RA, Casman SJ, Malyankar UM;
PI

PI Tchernev VT, Vernet CAM, Spyrek KA, Shenoy SG, Alsbrook JP;
PI

PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangoli EA, Boldog FI;
PI

PI Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;
XX

XX WPI; 2002-547774/58.
DR

DR N-PSDB; ABT05457.
XX

XX Novel isolated polypeptide, designated NOVX, useful for treating or
PT

PT preventing cancer, diabetes, obesity, dyslipidaemia, anorexia, and
PT

PT metabolic, neurodegenerative, immune and hematopoietic disorders.
XX

XX Claim 1; Page 44; 42ipp; English.
XX

XX The invention relates to an isolated polypeptide, designated NOVX,
CC

CC syndrome associated with a human disease, preferably a NOVX-associated
CC disorder, or for treating or preventing a NOVX-associated disorder in a
CC subject, preferably human. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are also useful
CC for treating or preventing metabolic disorders, diabetes, obesity,
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC disease, Parkinson's disorder, immune disorders, haematopoietic
CC disorders, and various dyslipidaemias, metabolic disturbances associated
CC with obesity, the metabolic syndrome X, wasting disorders associated with
CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This sequence
CC represents one of the isolated NOVX proteins of the invention
XX
XX Sequence 767 AA;
QY

Query Match 90.9%; Score 440; DB 5; Length 767;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTVLPHRLTEKRPDRINICIAQLKDLLPEHLKLTTLGHEKAVVLELTKHLKALTAL 104
Db 328 DTVLPHRLTEKRPDRINICIAQLKDLLPEHLKLTTLGHEKAVVLELTKHLKALTAL 387
QY 105 TEQHQKIIALQNGERSLKSPIQSDLDFAHSGFTCAKEVLQYLSRFSWTPREPCVQL 164
Db 388 TEQHQKIIALQNGERSLKSPIQSDLDFAHSGFTCAKEVLQYLSRFSWTPREPCVQL 447
QY 165 INHLHAVATQFLPTPOLLTQQVFLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 224
Db 448 INHLHAVATQFLPTPOLLTQQVFLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPVIQ 507
QY 225 RTQPSAELAAENDTDTDSGYGGEAEARPDPREKKGAGASRVTIKQEPGGSDSPAPKMKL 284
Db 508 RTQPSAELAAENDTDTDSGYGGEAEARPDPREKKGAGASRVTIKQEPGGSDSPAPKMKL 567
QY 285 DSRGGSGGGPGGAAAAAALLGPDPAALLLRPDAALLSLVAFGGGGGAPFPQAAA 344
Db 568 DSRGGSGGGPGGAAAAAALLGPDPAALLLRPDAALLSLVAFGGGGGAPFPQAAA 627
QY 345 AAAPFCFLPFCFLSPSAAAAYVQFLDKSGLEKLYPAAAAAPFLLYPGIPAAAAAAA 404
Db 628 AAAPFCFLPFCFLSPSAAAAYVQFLDKSGLEKLYPAAAAAPFLLYPGIPAAAAAAA 687
QY 405 AAAAAAAPPCLSSVLSPPPEKAGAAAATLTPHEVAPLGAHPHQPHQTHLPAGPRE 464
Db 688 AAAAAAAPPCLSSVLSPPPEKAGAAAATLTPHEVAPLGAHPHQPHQTHLPAGPRE 747
QY 465 PGNPSSAQEDPSQPGKEAP 484
Db 748 PGNPSSAQEDPSQPGKEAP 767

RESULT 10
AAU16188

ID AAU16188 standard; protein; 165 AA.
XX

AC AAU16188;
XX

XX 07-NOV-2001 (first entry)
DT

XX Human novel secreted protein, Seq ID 1141.
XX

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW

KW cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW

KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220953P.
PR 26-JUL-2000; 2000US-0220954P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 13-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234978P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488783/53.
N-PSDB; AAS26175.
New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.
Claim 11; SEQ ID NO 1141; 980pp; English.
The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
pathological condition. Antibodies to the proteins can also be used in
alleviating symptoms associated with the disorders and in diagnostic
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
(ELISA). Disorders which are diagnosed or treated include autoimmune
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,
nervous system disorders e.g. Alzheimer's disease, infections caused by
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
and many other disorders listed in the specification. The polypeptides
can also be used to aid wound healing and epithelial cell proliferation,
to prevent skin aging due to sunburn, to maintain organs before
transplantation, for supporting cell culture of primary tissues, to
regenerate tissues and in chemotaxis. The polypeptides can also be used
as a food additive or preservative to increase or decrease storage
capabilities, fat content, lipid, protein, carbohydrate, vitamins,
minerals, cofactors and other nutritional components. The present
sequence represents a novel secreted protein of the invention. Note: The
sequence data for this patent did not form part of the printed
Query Match 32.9%; Score 159; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. NO. 1.4e-132;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTVKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLTGLHLEKAVVLELTLLKHLKALTAL 104
Db 7 DTVKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLTGLHLEKAVVLELTLLKHLKALTAL 66
QY 105 TEOQHOKIIALONGERSLKSPTQSDLDAPHSGFQTCACEVLYLSRFSWTPREPRCQVL 164
Db 67 TEOQHOKIIALONGERSLKSPTQSDLDAPHSGFQTCACEVLYLSRFSWTPREPRCQVL 126
QY 165 INHLHVAVFQFLPTPQLLTQQVPLSKGTGCPAAGSAAA 203
Db 127 INHLHVAVFQFLPTPQLLTQQVPLSKGTGCPAAGSAAA 165
RESULT 11
ABU55257
ID ABU55257 standard; protein; 165 AA.
XX AC ABU55257;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #344.
XX KW Human; neural disorder; immune system disorder; renal disorder;

KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX Homo sapiens.
XX OS
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225475P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-147444/14.

DR N-PSDB; ABX73516.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX Claim 11; SEQ ID NO 1141; 402pp; English.
PS The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular polyps and sinusitis), respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 165 AA;

Query Match 32.9%; Score 159; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.4e-132;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLETLKHLKALTAL 104
Db 7 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLETLKHLKALTAL 66

Qy 105 TEQHQKIALLQNGERSLSKSPISQSDLDAPHSGFQTCAKEVLYLSRFESWTPEPRVCQL 164
Db 67 TEQHQKIALLQNGERSLSKSPISQSDLDAPHSGFQTCAKEVLYLSRFESWTPEPRVCQL 126

Qy 165 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAA 203
Db 127 INHLHAVATQFLPTPOLLTQQVPLSKGTGAPSAAGSAAA 165

RESULT 12
ABO14646
ID ABO14646 standard; protein; 203 AA.
AC ABO14646;
XX
XX 25-AUG-2003 (first entry)
DT Novel human protein #19.
XX
XX Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection.
XX
OS Homo sapiens.
XX
XX WC2003023002-A2.
PN
XX 20-MAR-2003.
PD
XX 09-SEP-2002; 2002WO-US028539.
PF
XX
XX

PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 17-SEP-2001; 2001US-0322817P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX (CURA-) CURAGEN CORP.
XX
XX SPYTEK KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shenoy SG, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
DR WPI: 2003-313242/30.
DR N-PSDB; ACD19339.
XX
XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
XX Claim 1; Page 126; 586pp; English.
XX
XX The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 203 AA;

Query Match 22.3%; Score 108; DB 6; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.1e-87;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLETLKHLKALTAL 104
Db 45 DTYKLPRLIEKRRDRINECIAQLKDLLPEHLKLTTLGLHLEKAVVLETLKHLKALTAL 104

Qy 105 TEQHQKIALLQNGERSLSKSPISQSDLDAPHSGFQTCAKEVLYLSRF 152
Db 105 TEQHQKIALLQNGERSLSKSPISQSDLDAPHSGFQTCAKEVLYLSRF 152

RESULT 13
AAU16603
ID AAU16603 standard; protein; 150 AA.

XX AAU16603;
AC 07-NOV-2001 (first entry)
XX Human novel secreted protein, Seq ID 1556.
DT
XX
DE
XX
KW Human; immunosuppressive; antiarthritic; anirrhematic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX
PN WO20015322-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180828P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI N-PSDB; AAS26590.
XX
DR WPI; 2001-488783/53.
DR N-PSDB; AAS26590.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1556; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathologic condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 22.1%; Score 107; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 LGHLEKAVVLELTALKHKTALTECHOKITLALONGERSLKSPIQSOLDLAFHSGFQCA 141
Db 44 LGHLEKAVVLELTALKHKTALTECHOKITLALONGERSLKSPIQSOLDLAFHSGFQCA 103
QY 142 KEVLYLSRFESWTPEPRCVQLINHLHAVAQTFLTPQLLTQQVPL 188
Db 104 KEVLYLSRFESWTPEPRCVQLINHLHAVAQTFLTPQLLTQQVPL 150
RESULT 14
ABU55672
ID ABU55672 standard; protein; 150 AA.
XX
XX AC ABU55672;
XX PA

DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #759.
XX
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
XX US2002132753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216980P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225677P.
XX 14-AUG-2000; 2000US-0225688P.
XX 14-AUG-2000; 2000US-0225707P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 22-AUG-2000; 2000US-0226888P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 08-SEP-2000; 2000US-0231413P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 27-SEP-2000; 2000US-0235834P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 29-SEP-2000; 2000US-0236370P.
XX 02-OCT-2000; 2000US-0236802P.
XX 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
XX 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
XX 13-OCT-2000; 2000US-0239935P.
XX 20-OCT-2000; 2000US-0240960P.
XX 20-OCT-2000; 2000US-0241785P.
XX 20-OCT-2000; 2000US-0241809P.
XX 17-NOV-2000; 2000US-0244617P.
XX 17-NOV-2000; 2000US-0249399P.
XX 08-DEC-2000; 2000US-0251856P.
XX 08-DEC-2000; 2000US-0251868P.
XX 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX
XX PA

PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.
DR N-PSDB; ABX73931.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1556; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 150 AA;
Query Match 22.1%; Score 107; DB 6; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 LGHLEKAVVLETLKHLKALTALTEQCHOKIIALONGERSLKSPICQSLDAPHSGFQCA 141
Db 44 LGHLEKAVVLETLKHLKALTALTEQCHOKIIALONGERSLKSPICQSLDAPHSGFQCA 103
QY 142 KEVLQYLSRFESWTPREPCVQLNHLHVAATQFLTPOLLTQQVPL 188
Db 104 KEVLQYLSRFESWTPREPCVQLNHLHVAATQFLTPOLLTQQVPL 150
RESULT 15
AAB70694
ID AAB70694 standard; protein; 410 AA.
XX
AC AAB70694;
XX
DT 17-MAY-2001 (first entry)
XX
DE Mouse DEC2a protein sequence SEQ ID NO:14.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX
OS Mus musculus.
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
XX
DR WPI; 2001-202935/20.

DR N-PSDB; AAF74777.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
PS Claim 2; Page 77-80; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed mouse DEC2a protein, as given in the
CC present invention
XX
SQ Sequence 410 AA;
Query Match 16.7%; Score 81; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 5.3e-63;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 104
Db 43 DTYKLPRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLETLKHLKALTAL 102
QY 105 TEQCHOKIIALONGERSLKSP 125
Db 103 TEQCHOKIIALONGERSLKSP 123
Search completed: December 15, 2004, 20:10:58
Job time : 164 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2498	99.2	482	2	JC7583	basic helix-loop-h
2	1634	64.9	410	2	JC7584	basic helix-loop-h
3	669	26.6	412	2	JCS547	basic helix-loop-h
4	209.5	8.3	460	2	T33110	hypothetical prote
5	188	7.5	218	2	S29712	segmentation prote
6	187	7.4	280	2	A53027	transcription fact
7	183.5	7.3	281	2	S36748	transcription fact
8	182.5	7.3	435	2	A46231	hypothetical helix p
9	182	7.2	581	2	T22341	hypothetical prote
10	181	7.2	465	2	G03738	FRAC-4 - human
11	180	7.2	418	2	T15142	hypothetical prote
12	179.5	7.1	676	1	EDBE23	immediate-early pr
13	177	7.0	282	2	A53336	transcription fact
14	176	7.0	846	2	S53418	GTP-binding regula
15	175.5	7.0	627	2	A41112	spidroin 2, dragli
16	172.5	6.9	416	1	SKL1AG	dermal gland prote
17	171.5	6.8	420	2	A49642	transcription fact
18	171.5	6.8	676	1	EDBE22	immediate-early pr
19	171.5	6.8	801	2	T29018	hypothetical prote
20	171.5	6.8	3534	2	T45667	tegument protein 2
21	170.5	6.8	825	1	EDBEXD	immediate-early pr
22	170	6.8	901	2	A49227	sialidase - Actino
23	169	6.7	702	2	G01840	T-box protein 2 -
24	169	6.7	862	2	T48289	hypothetical prote
25	168.5	6.7	416	2	SL2541	evx1 protein - mur
26	167.5	6.7	903	2	T00705	N-chimerin homolo
27	167	6.6	780	2	T00366	hypothetical prote
28	167	6.6	2639	2	T31328	fibrin - Chinese
29	164.5	6.5	1334	2	T50568	probable multi-dom

QY 361 AAAYVQPFLDKSGLEKYLYPAAAAA
FPLLYPGIPAPAAAAAAAFPCLVSV 42

C:Genetics:

A:Gene: CESP:C18H7.3

A:Map position: 4

A:Introns: 84/1

C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 8.3%; Score 209.5; DB 2; Length 460;

Best Local Similarity 31.5%; Pred. No. 2.9e-05; Indels 81; Gaps 14;

Matches 99; Conservative 15; Mismatches 119; Indels 81; Gaps 14;

QY 193 GAPSAAGS---AAAPCLERAGOKLEPLAYCVPIQTOPSABLAENDTDTDSGYGGA 249

DB 186 GRKPAAGSDGAGAGCTKTNPAGP-----GPAGPPGPPGNGRAGAGSGPGPGP 237

QY 250 APDPREKKGK--AGASRVTIKEPPGEDSPAPKRMKLDNRGGSGGGPGGAA----- 300

DB 238 AGPPGNGKDGEGA-----PGKGDG-----GADGAPGTDAAYCPCPPR 277

QY 301 -AAAAALLGPDAAA---AALLRPDAALLSLVAFGGGGGAPFPAPAAAAPFCLPFCFL 356

DB 278 SAALGAGGAGPAGAPAAPEAAAPAAPEAAPEAAEGAGGA---SPAGGAAP---DAAA 331

QY 357 SPFAAAYVQPFIDKSGLEKYLPA--AAAPFLYPCIPAPAAAAA-----AAAF 414

DB 332 APEAAPAAPEAAEGAGGAPAGAPDAAAAPAAPEAAPEAAEGAGGAPAGAAA- 390

QY 415 PCLSSVLSPPPPEKAGAAAATLLPHEVAPL--GAPHPQHGHGTHLFPFAGPREPGN--PES 470

DB 391 -----PDAAAPEAAAPAAPEAAEG-----GGGAPAGAAPEA 426

QY 471 SAQEDPSQPGKEAP 484

DB 427 AAAAPGAGGGEAP 440

RESULT 5

S29712

segmentation protein hairy - red flour beetle

C:Species: Tribolium castaneum (red flour beetle)

C:Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 09-Jul-2004

A:Reference number: S29712

A:Accession: S29712

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-218 <SOM>

A:Cross-references: UNIPROT:Q8T306

Query Match

Best Local Similarity 30.1%; Pred. No. 0.00024;

Matches 72; Conservative 28; Mismatches 91; Indels 48; Gaps 10;

QY 53 LIEKRRDRINECIAQLKDLPEHLKLTTLGH--LEKAVVLETLKHLKALTALTEQHQ 110

DB 5 IMEKRRARRINNSLEKTLILDAMKKOPARHSKLEKADILEMTVKHLQNL-----ORQ 58

QY 111 KITALQNGERSLKSPIQSDLDAPHSGFQTCAKEVLQYLSRFSWTPRPRCVQLINHLHA 170

DB 59 QAAMWQTPDPSVVS-----KFRAGFSCASEV-----GRFGLDPVVKR--RLIQHLAS 105

QY 171 VATQFLPTQ-----LITQOVPLSKGTGAP-----SAAGSAAAPCLERAGOKLEPLAY 218

DB 106 CLNQGQKEPQVQVIVPEVAPNNILNGTGVQLVPIRLANGDIA---LVLPTQGASPLFL 162

QY 219 CVPVIQTOPSABLAENDTDTDSGYGGAEARPDREKKGAGASRVTIKEPPGEDSP 277

DB 163 LVPIPORTASTASSA-----SNYSPQSPEPESVR-----PLSLVVRREETPEERK 209

RESULT 6

A53027

transcription factor HES-1 - human

N:Alternate names: hairy protein homolog

C:Species: Homo sapiens (man)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

A:Accession: A53027

R:Pedfer, J.N.; Li, L.; Jan, L.Y.; Jan, Y.N.

Genomics 20, 56-61, 1994

A:Title: Genomic cloning and chromosomal localization of HRY, the human homolog to the D.

A:Reference number: A53027; MUID:94292187; PMID:8020957

A:Accession: A53027

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <PED>

A:Cross-references: UNIPROT:Q14469; GB:L19314; NID:g436999; PIDN:AAA65220.1; PID:g780128

C:Genetics:

A:Gene: GDB:HRY

A:Cross-references: GDB:217077; OMIM:139605

A:Map position: 3q28-3q29

C:Keywords: immediate-early protein

Query Match 7.4%; Score 187; DB 2; Length 280;

Best Local Similarity 27.5%; Pred. No. 0.00036;

Matches 86; Conservative 25; Mismatches 88; Indels 114; Gaps 14;

QY 31 KPFRSMKRDCTKVSDTYKLPRLHLEKKRRDRINECIAQLKDLPEHLKLTTLGH--LEKA 88

DB 27 KPFTA-----SEHKSSKPINEKRRRARRINESQLKTLILDALKDSSRHSKLEKA 78

QY 89 VVLETLKHLKALTALTEQHQKIIALQNGERSLKSPIQSDLDAPHSGFQTCAKEVLQYL 148

DB 79 DILEMTVKHLRNL-----QRAQMTAAL-----STDPSVLGKYRAGFSECKNNEVTRFL 125

QY 149 SRFESWTPRPRCVQLINHLHAVATQF-----LPTPQLLTQQVPLSKGTGAPSAAGSA 201

DB 126 STCEG-VNTEVR-TRLGLHLANCMTQINAMTYCQPHPALQAPPPP-PPGPGGPGHAPFA 182

QY 202 AAPCLERAGOKLEPLAYCVPIQTOPSABLAENDTDTDSGYGGAEARPDREKKGAG 261

DB 183 PPP-----PL---VFI-----

QY 262 ASRVTIKEPPGEDSPAPKRMKLDNRGGSGGGGPG--GGAAAAAALLG-----PDAAA 314

DB 191 -----PGGAAPP-----GGAPCKLGSQAGAAKVFQGVVVPADPGQ 228

QY 315 AALLRPDAALLSS 327

DB 229 FAFLLPNGAPHS 241

RESULT 7

S36748

transcription factor HES-1 - rat

N:Alternate names: hairy-like gene RH1; helix-loop-helix protein

C:Species: Rattus norvegicus (Norway rat)

C:Date: 09-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

A:Accession: S36748; A48106

R:Shasai, Y.; Kageyama, R.; Tagawa, Y.; Shigemoto, R.; Nakanishi, S.

Genes Dev. 6, 2620-2634, 1992

A:Title: Two mammalian helix-loop-helix factors structurally related to Drosophila hairy

A:Reference number: S36748; MUID:94040724; PMID:1340473

A:Accession: S36748

A:Molecule type: mRNA

A:Residues: 1-281 <SAS>

A:Cross-references: UNIPROT:Q04666; EMBL:D13417; NID:g220760; PIDN:BAA02682.1; PID:d1003

R:Pedfer, J.N.; Jan, L.Y.; Jan, Y.N.

Mol. Cell. Biol. 13, 105-113, 1993

A:Title: A rat gene with sequence homology to the Drosophila gene hairy is rapidly induc

A:Reference number: A48106; MUID:93109293; PMID:8417318

A:Accession: A48106

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-85,'N',87-92,'W',94-281 <PED>

A;Cross-references: GB:L04527; NID:G204554; PIDN:AAA41307.1; PID:G204555
A;Note: sequence extracted from NCBI backbone (NCBIN:121122, NCBI:P:121123)
C;Keywords: DNA binding; transcription regulation

Query Match 7.3%; Score 183.5; DB 2; Length 281;
Best Local Similarity 27.0%; Pred. No. 0.00058;
Matches 84; Conservative 25; Mismatches 93; Indels 109; Gaps 12;

Qy 31 KPKSRMKRDTKVSQDYKPLPHRLIEKKRRDRINECIAQLKDLLPEHLKTLTGH--LEKA 88
Db 27 KPNTA-----SEHRKSKTMEKRRRINELSLQKTLILDALKDSSRRHLEKA 78
Qy 89 VVLELTLELTKALTALTEQHQKIIALQNGERSLSKPSQSDIDAFHSQFQTCACEVLOYL 148
Db 79 DILEMTVKELRNL-----QRAQWTAAL-----STDPVLGKYRAGFSECMNEVTRFL 125
Qy 149 SRPESWTPREPRVCVQLINHLHVAIQF-----LFTPOLLTQOVPLSKGTGAPSAAGA 201
Db 126 STCEG-VNTEVR-TRLLGHLANCMQINAWTVPGQAHPALQAPPPPPSPGPGQHPAPFA 183
Qy 202 AAPCLERAGKLEPLAYCVPIQRTQPSAEALAAENDTDTDSYGGEAEARPREKKGAG 261
Db 184 PPP-----PL--VPI-----191
Qy 262 ASRVTIKQPPGSDSPAPKMKLDSRGSGGGGPGGAAAAAALLG-----PDPAAGAAA 316
Db 192 -----PGGAAPP-----GSAFCKLGSQAGEAAKVFQGVVPAAPDGOFA 231
Qy 317 LLRPDAALLSS 327
Db 232 FLIPNGAFHS 242

RESULT 8
A46231
helix-loop-helix protein Dpn - fruit fly (*Drosophila* sp.)
C;Species: *Drosophila* sp.
C;Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 10-Sep-1997
C;Accession: A46231
Genes Dev. 6, 2137-2151, 1992
R;Bier, E.; Vaessin, H.; Younger-Shepherd, S.; Jan, L.Y.; Jan, Y.N.
A;Title: deadpan, an essential pan-neuronal gene in *Drosophila*, encodes a helix-loop-helix
A;Reference number: A46231; MUID: 93051333; PMID: 1427077
A;Accession: A46231
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-435 <BIE>
A;Cross-references: GB:S48025; NID:G259771; PID:G259772
A;Note: sequence extracted from NCBI backbone (NCBIN:117859, NCBI:P:117860)

Query Match 7.3%; Score 182.5; DB 2; Length 435;
Best Local Similarity 21.7%; Pred. No. 0.0011;
Matches 97; Conservative 53; Mismatches 146; Indels 151; Gaps 17;

Qy 24 YSSLYMCKPKRSMKRDITKVSQDYKPLPHRLIEKKRRDRINECIAQLKDLLPEHLKTLTGL 83
Db 19 YSDSYSGNGRSMNPGLSK-AELRKTNKPIEKKRRRINELHCLNELSLILEAMKDPAR 77
Qy 84 H--LEKAVVLELTLELTKALTALTEQHQKIIALQNGERSLSKPSQSD---LDAFHSQFQ 138
Db 78 HTKLEKADILEMTVKELQSV-----QROQ-----LNMAIQSDPSVVQVKFKTGTV 121
Qy 139 TCAKEVLYQLSRPES-----WTPREPRVCVQLINHL-----HAVAQFQ 175
Db 122 ECAEEVNRVYQMDGIDTGVQRLSAHLNQCANLEIQGSMNSFNFGYRGGLFPATAYTA 181
Qy 176 LFTPOLLTQOVPLSKGTGAPSAAGSAAPCLERAGKLEPLAYCVPIQRTQPSAEALAE 235
Db 182 APTPLPFLPQDLNNNSRTES-----SAPAIQMGGLQIIP-----SRLPSGEFA-- 225
Qy 236 NDTDTDSYGGEAEARPREKKGAGASRVTIKQEPGSDSPAPKMKLDSRGSGGGGP 295
Db 226 -----LIMPTGSAAPP-----GPFANP 244

Qy 296 GGGAAAAAALLGPDPAAGAAALLRP-----DAALLSSLVAFGGGGAPFPQPAAGAAAPPCL 351
Db 245 GSAAGVAA-----GTASAAALASIANPHLNDYTOSEFMSAFS-----KPVNTSYFANL 292
Qy 352 PFCFLSPSAAAYVQPPFLDKSGLEKLYPAAAPFPFLLYPGIPAPAAAAAAGAAAAA 411
Db 293 P-----ENIHTLPGOTQLPVKNSTSPPLSPSSISSHCEESR 330
Qy 412 AAPCLSSVLS-----PPEKA 428
Db 331 AASFTV-DVLSKHSFAGVSTPTPTSA 356

RESULT 9
T22341
hypothetical protein F47B8.5 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22341
R;Berks, M.; Murray, A.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19551
A;Accession: T22341
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-581 <WIL>
A;Cross-references: UNIPROT:Q20517; EMBL:Z77662; PIDN:CAB01192.1; GSPDB:GN00023; CBSP:F4;
A;Experimental source: clone F47B8
C;Genetics:
A;Gene: CBSP:F47B8.5
A;Map position: 5
A;Introns: 37/3; 82/3; 122/3; 148/3; 194/1; 217/3; 298/2; 368/1

Query Match 7.2%; Score 182; DB 2; Length 581;
Best Local Similarity 31.9%; Pred. No. 0.0016;
Matches 72; Conservative 5; Mismatches 77; Indels 72; Gaps 7;

Qy 286 SRGGSGGGGPGGAAAAAALLGPDPAAGAAALLRPDAALLSSLVAFGGGGAPFPQPAAG 345
Db 389 SSGGGGGYSGGGGGGGAAPPSPADAAAPAPP-----APEPVFAPA 432
Qy 346 AAPFLCPLFCFLSPSAAAYVQPPFLDKSGLEKLYPAAAPAP-----FPL----- 389
Db 433 PAP-----EAPVAPSADAGY-----AAAAAPAGGSGYPAKKRRVAEDYAE 474
Qy 390 -----LYPGIPAPAAAAAAGAAAAAAPPCLSSVLSPPPEKAGAAATLLP-----H 438
Db 475 GEAP 534
Qy 439 EVAPLGAHPHQHGRTHLPFAGPREPGNPESQAQEDRSPQKPEAP 484
Db 535 EVAPAPAPAP-----AAPESDASGYSSGAAAPAGCGGSGYP 569

RESULT 10
G02738
FREC-4 - human
C;Species: *Homo sapiens* (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C;Accession: G02738
R;Ehrback, S.
submitted to the EMBL Data Library, June 1996
A;Reference number: H01646
A;Accession: G02738
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-465 <ENE>
A;Cross-references: UNIPROT:Q16676; EMBL:U59832; NID:gl399238; PID:gl399239
C;Genetics:
A;Gene: FREAC-4
C;Superfamily: fork head DNA-binding domain homology
P:125-216/Domain: fork head DNA-binding domain homology <FHD>

Query Match 7.3%; Score 181; DB 2; Length 465;
Best Local Similarity 29.5%; Pred. No. 0.0014;
Matches 100; Conservative 21; Mismatches 104; Indels 114; Gaps 19;

```
QY 235 ENDTDTDSGVGG---EAAAPDPREKKGAGASRVTKOE-----PCGEDSPAP- 279
Db 31 EEEEDDEGGGGPRLVAPQRRRRRSVAGEDLEEEEDDDILLAPPAGGSPAPP 90
QY 280 ---KRWKLDSRGSGGGGPGGGAAGAAA---LLGPDPAAGAAAALL-----RPDAAL-LS 326
Db 91 GPAPAAAGAGAGGGGGGGGAGGGSAGSAGAKNPLVKP-PYSYIALITMAILQSPKKLTL 149
QY 327 SLVAFGGGGGAPPO-----PA---AAAPFCIPFCF-----LSPSAA 361
Db 150 EICEFISG---RPYYREKFPWQNRHNLNLUNCDFVKIPREPFGNGKGYWTLDPESA 206
QY 362 AAY-----VQPFLL--DKSGLEKLYL-----PAAAAAPFFLLYPGIP 395
Db 207 DMFNGDSFLRRKRFRKQPLLPNAAAESLLRGAGAGGAGDPAAAAALLFPAPPPPP 266
QY 396 -----APAAAAAAGAAAAAFAFPCLSVLSPPEKAGA-AAATL 435
Db 267 HAYGYGYGCGYGLQFPYAPPSALFAAAAAAFAFHPSPPPPPPPHGAAAEALARTA 326
QY 436 LPHEVAPLGA--PHPQHPHGRTHLPEAGPREPGNPRESSA 472
Db 327 FGYPHPPLGAALDGP-----LP-ASAKAGGPGGASA 356
```

RESULT 11
T15142
hypothetical protein T28F2.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15142
R:Madson, C.; Fronick, B.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of *C. elegans* cosmid T28F2.
A:Reference number: Z18300
A:Accession: T15142
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-418 <MAD>
A:A:Cross-references: UNIPROT:O01662; EMBL:AF000198; NID:G2047345; PID:G2047346; PIDN:AAAB5
A:Experimental source: strain Bristol P2, clone T28F2
C:Genetics:
A:Gene: CESP:T28F2.6
A:Map position: 1
A:Introns: 49/3
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match	7.2%	Score 180;	DB 2;	Length 418;
Best Local Similarity	32.0%;	Pred. No. 0.0014;		
Matches	93;	Conservative 6;	Mismatches 120;	Indels 72; Gaps 14;
187	QY	PLSKGTGAPSAAGSAAPCLERAG-OKLBLEYACVPIQRTQPSAEALAAENDTDTDSGVG	245	
173	DB	PGPDGNAGPAGPAGAGPGEAGNYAEPGPAGPAGPDPGDQCAP-----GPDGQFGAG	227	
246	QY	GEAEARPDEKKGAGASRVTIKQEPFGBEDSPAPKMKLDSRGGSGGGSPGGGAAAAAAA	305	
228	DB	GTTSTNQPPGP-AG-----PPFGPAGPAGEDAYAQPSPAGTGPFGPPGPKDGEA-	276	
306	QY	LLGDPDAAAA---ALLRPDAALLSLVAFGGGGGAPFPQPAAAAAAPFCLFCFLSFSAA	361	
277	DB	-GPDGFPAGPGTGDGAPGDDAAVCPCPRTLGAAGYPEGDDAAARP-----A	322	
362	QY	AAVQPFDLKSGLEKLYPAAAAAPFLLYPGTAPAAAAAAA-----AAAAAAAFPCU	417	
323	DB	GGYD-----GGAGAAPEAAAPAAAAAPG-----APAPAAAAAPAGGYGGGAAAAAA	369	
418	QY	SSVLSPPPPKAGAAAAATLLPHEVAPLGCAPHQPHGRHTLHPFAGPPGPNP	468	

Db 370 -----PPPPAAAAAEPAP---APAAAPPP-----APAAAG-----GSP 400

RESULT 12

EDBB23

immediate-early protein IER2.9 - bovine herpesvirus 1 (strain Jura)

N/Alternate names: early protein ER2.6; p135 protein

C/Species: bovine herpesvirus 1

C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004

C/Accession: B38209

R/Wirth, U. V.; Eraefel, C.; Vogt, B.; Vlcek, C.; Paces, V.; Schwyzer, M.

J. Virol. 66, 2763-2772, 1992

A/Rtitle: Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1 are 3' coterm

A/Reference number: A38209; MUID:92219360; PMID:1313901

A/Accession: B38209

A/Molecule type: DNA

A/Residues: 1-676 <WR>

A/Cross-references: UNIPROT:P29128; GB:M84465; NID:G330769; PIDN:AAA46062.1; PID:G330770

C/Superfamily: herpesvirus immediate-early protein IER1; RING finger homology

C/Keywords: DNA binding; immediate-early protein; transcription regulation; zinc finger

F/9-57/Domain: RING finger homology <RING>

F/13-51/Region: zinc finger C3HC4 motif

F/284-331/Region: acidic

```

Query Match      71.1%; Score 179.5; DB 1; Length 676;
Best Local Similarity 28.5%; Pred. No. 0.0026;
Matches      92; Conservative 35; Mismatches 125; Indels 71; Gaps 15;

Qy 171 VATQFLPTTQLLTQQVPLSGKGTGASAGSAAAPCLERA-GOKLELAYCVPVIOQTQS 229
Db 352 VSTRGQTFAV--QPAPRSLARRPCGAAAAVAPSSRSRGRDRP---LGAAPAAFA 406
Qy 230 AELAAENDTDTDSYG-----GE-AEARPDREXKGAGASRVTIKQBPFGEDSPAPKR 281
Db 407 AQARACSPFEREGRCAGLGVAGETAGCWGASGSGRGERRAELLGEAGPPVQVARRRR 466
Qy 282 MKLDSRGSGGGPGGGAAAAAALLGPPDAAAAALLRPDAALLSLVAFGGGGA--PF 339
Db 467 TELDR-----APTPAPAPA---PAPAPISTVIDLTANAPARPAD 502
Qy 340 PQPAAAAAPFCLPFCELSAAAAAYOPFLDKSGLEKILYPAAAAAPFLLVPGTPAPAA 399
Db 503 PAPAAAAPGPASAGAQIGTPAAAAAVT-----AAAAAP-SVARSSAPSPAV 546
Qy 400 AAAAAAASAAAAAAPPCLSGSVLSP-----PEKAGA-AAATLLPHEVAPLGAHPQHPGHT 455
Db 547 TAAATSTAAAIITRAPTPSPAGAPAADPRRACAPALAGAAAEVGRGNP-----GRE 600
Qy 456 HLFFA----GPREPGNPSSAQE 474
Db 601 BRPASAMARGDLDPG-PSSACK 622

```

```

RESULT 13
A53336
transcription factor HES-1 - mouse
N;Alternate names: hairy protein homolog
C;Species: Mus musculus (house mouse)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A53336
R;Rakabayashi, K.; Sasai, Y.; Sakai, Y.; Watanabe, T.; Nakanishi, S.; Kageyama, R.
J. Biol. Chem. 269, 5150-5156, 1994
A;Title: Structure, chromosomal locus, and promoter analysis of the gene encoding the mo-
A;Reference number: A53336, MUID:94148977; PMID:7906273
A;Accession: A53336
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-282 <PAK>
A;Cross-references: UNIPROT:P35428; GB:D16464; NID:G475013; PIDN:BA03931.1; PID:di100444
C;Genetics:
A;Introns: 36/3; 68/3; 98/1
C;Keywords: DNA binding; transcription regulation

```

```
Query Match      7.0%; Score 177; DB 2; Length 282;
Best Local Similarity 26.7%; Pred. No. 0.0014;
Matches 83; Conservative 25; Mismatches 95; Indels 108; Gaps 12;

QY 31 KPESMKRDDTKVSDYKYLPHALIEKKRRDRINECIAOLKDLPEHLKLTTLGH--LEKA 88
DB 27 KPETA-----SEHRKSKPINEKRRRINESLSQKTLILDAIKOSSRHHSKLEKA 78
QY 89 VVLELTAKHLKALTALTEQOQHOKIITALONGERSLSPIQSDLDAPHSGFQTCACEVLQYL 148
DB 79 DILEMTVKHLRNL-----QRAQMTAL-----STDPSVLGKYRAGSECMNEVTRFL 125
QY 149 SRPESWTPEPRCVQLINHLHVAATQF-----LPTPQLLTQQVPLSKGTGAPSAAGA 201
DB 126 STCEG-VNTEVR-TRLGLHLANCMQTINAMTYPGQAHPALQAPPPPPSPGAPQHPFA 183
QY 202 AAPCLERAGQKLEPLAYCVPIQRTQPSAEALAEENDTDTDSGYGGEAEARPDRKKGAG 261
DB 184 PPP-----PPL---VPI----- 192
QY 262 ASRVTIKQPPGSDSPAPKRMKLDNRGGSGGPGGGAALIG-----PDPAALAAA 316
DB 193 -----PGGAAPP-----GSAPCKLGSQAGEAAKVFGGFQVVPAPDQGOFA 232
QY 317 LLRPDAALLSS 327
DB 233 FLIPNGAFHS 243

RESULT 14
S52418
GTP-binding regulatory protein Gs alpha-XL chain - rat
N:Alternate names: G protein XL alpha-s
C:Species: Rattus norvegicus (Norway rat)
C>Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C:Accession: S52418
R:Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 372, 804-809, 1994
A:Title: XL-alpha-s is a new type of G protein.
A:Reference number: S52418; MUID:95089824; PMID:7997272
A:Accession: S52418
A:Molecule type: mRNA
A:Residues: 1-846 <KEH>
A:Cross-references: EMBL:X80407; NID:G642267; PIDN:CAA58866.1; PID:G642268
R:Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 375, 253, 1995
A:Title: Correction: Xlalphas is a new type of G protein.
A:Contents: annotation; assignment of start_codon
A:Note: experimental data from this paper suggest that the translation is initiated at P
C:Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F:132-846/Product: GTP-binding regulatory protein Gs alpha-XL chain #status experimental
F:499-506/Region: nucleotide-binding motif A (P-loop)
F:744-747/Region: GTP-binding NKXD motif

Query Match      7.0%; Score 176; DB 2; Length 846;
Best Local Similarity 25.1%; Pred. No. 0.0054;
Matches 106; Conservative 37; Mismatches 133; Indels 146; Gaps 17;

QY 105 TEQOQHOKIITALONGERSLSKP-----TQSDLDAPHSGFQICAKEVLQYLSRFESW 154
DB 69 TTAEAKVPSLEKGEQSPQETVHTKPAVRESGTDSSKADPDSATHAVLQI----- 121
QY 155 TPPEPRCVQLINHLHVAATQFLPTPQLLTQQVPLSKGTG-----APS----- 196
DB 122 GPEVGGV-----PTWPTDLPASSEDAGPDVRAEPDGGTATAPAESGD 165
QY 197 -----AAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAEALAEENDTDTDSGYGGEAEARP 252
DB 166 NREPAALAAAAPAEPAEFAA---EPAA-----EPAAEPAAE-----PAEAVP 203
QY 253 DREKKGAGASRVTIKQEPPEGSDPAPKRMKLDNRGGSGGPGGGAALIGLDPDA 312
```

```
DB 204 DTEASASGA--VPDQEEP-----AAAAATPAEPA 234
QY 313 AAAALL---RFDAAALLSSLVAFGGGGGAPFPAPAAAAAPCLPFCFLSPSAAAAYVQPEL 369
DB 235 ARAPVTPTEPATRAVPSASAHFAAGAVPGASAMSA---RAAARAAVAGPL- 285
QY 370 DKSLEKLYLPAAAAAPFLLYGIPAPAAAAA-----AAAFCLSLVLP 423
DB 286 -----VWGAASLSLATPAARASLPAARAAAAAARAAVAAGSASAPRAHLRPP 337
QY 424 PPEKAGAAAAATLLPHEVAPLGAHPH-----QHPHGRTHLPPFAGPRPFGNPESQAQEDPSQ 478
DB 338 SPEIQVADPT-----PPAPRPSANWPKYERGRSCCRYE-ASGICEIESSDESE 388
QY 479 PG 480
DB 389 EG 390

RESULT 15
A44112
spidroin 2, dragline silk fibroin - orb spider (Nephila clavipes) (fragment)
N:Alternate names: silk fibroin, dragline
C:Species: Nephila clavipes
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A44112; S27824
R:Hinman, M.B.; Lewis, R.V.
J. Biol. Chem. 267, 19320-19324, 1992
A:Title: Isolation of a clone encoding a second dragline silk fibroin. Nephila clavipes
A:Reference number: A44112; MUID:92406876; PMID:1527052
A:Accession: A44112
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-627 <HIN>
A:Cross-references: UNIPROT:P46804; GB:M92913; NID:G159713; PID:G159714
A:Note: sequence extracted from NCBI backbone (NCBI:P:113993)
R:Hinman, M.B.; Lewis, R.V.
submitted to the EMBL Data Library, May 1992
A:Description: Isolation of a clone encoding a second dragline silk fibroin: Nephila clava
A:Reference number: S27824
A:Accession: S27824
A:Molecule type: mRNA
A:Residues: 19-627 <H12>
A:Cross-references: EMBL:M92913

Query Match      7.0%; Score 175.5; DB 2; Length 627;
Best Local Similarity 25.9%; Pred. No. 0.0041;
Matches 109; Conservative 20; Mismatches 135; Indels 157; Gaps 18;

QY 184 QQVPLSKGTG--APSAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAEALAEENDTDT 241
DB 93 QQQPGGYGQQGPGSGPSGASAAA-----ASAAASAESQGGP 128
QY 242 SGYG-----GEAEARPDREKKGAG-ASRVTIKQEPPEGSDP---APKRMK 283
DB 129 GGYGPGQQGPGGYGPGQQGPGGPGGPGGSAASGPGGPGGPGGPGGPGGPGGPGG 188
QY 284 LDRSGGGSGG--GPGGGAALAAA-----LIGPDPA 314
DB 189 PGGYGPQQGPGSGPSGASAAAASGPGQPGGYGPGGPGGPGGPGGPGGPGGPGG 248
QY 315 AALL-----PDAALLSSLVAFGGGGGAPFPAPAAAAAPCLPFCFL 356
DB 249 AAAAGPQQGPGGYGPGQQGPGSGSAAAAAAGPGGYGPGGPGGPGGPGGPGGPGG 308
QY 357 SPASAAAAYVQFLDKSGLEKY-----LYPAAAP---FP 388
DB 309 GSAAAAAAGP--GQQLGGYGPGQQGPGGYGPGGPGGPGGPGGPGGPGGPGGPGG 366
QY 389 LLY-PGIPAPAAAAA-----FPCLSVLSPPEKAGAAA 433
DB 367 GGYGPGQQGPGSGPSASAAAAAAGPGGYGPGQPGGYGPGGPGGPGGPGGPGGPGG 426
```

QY 434 TLLPHEVADLG-APHPOHPGRTHLIP-FAGPREPGNPSS-----AQEDPSQP 479
Db 427 A-----AGFGGYPGQQGEGG--YAPGQQGPGSGSAAAAAAGPGGYGPAQQGSGP 479
QY 480 G 480
Db 480 G 480

Search completed: December 15, 2004, 19:56:33
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2004, 19:15:04 ; Search time 154 Seconds
(without alignments)
1127.435 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 2517
Sequence: 1 MDEGPHLOERQLLEHRDFI.....PQNPSSAQEDPSQPCKEAP 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2517	100.0	484	4	AAB70693 Human DEC
2	2498	99.2	482	4	AAB70692 Human DEC
3	2498	99.2	482	5	ABG96308 Human ova
4	2498	99.2	482	6	ABO14645 Novel hum
5	2498	99.2	482	6	ABR58470 Human BHL
6	2498	99.2	482	7	ADC31646 Human nov
7	2498	99.2	482	7	ADE34662 Human Pro
8	2498	99.2	482	8	ADN05014 Antipsori
9	2498	99.2	482	5	ABJ04644 Protein o
10	1634	64.9	410	4	AAB70694 Mouse DEC
11	1624	64.5	410	7	ADE34660 Rat Prote
12	905	36.0	203	6	ABO14646 Novel hum
13	835	33.2	185	4	AAU16188 Human nov
14	835	33.2	185	6	ABU55257 Human nov
15	750	29.8	150	4	AAU16603 Human nov
16	750	29.8	150	6	ABU55672 Human nov
17	685.5	27.2	404	5	ABBS7218 Mouse isc
18	685.5	27.2	411	7	ADD90602 Mouse BHL
19	685.5	27.2	411	8	ADP74022 Murine DE
20	682.5	27.1	411	7	ADE34664 Rat Prote
21	669	26.6	412	2	AAW08140 Human cyt
22	669	26.6	412	2	AAW99060 Human bas
23	669	26.6	412	3	AAV87958 Human CR8
24	669	26.6	412	4	AAB30758 Human she
25	669	26.6	412	5	ABP65203 Hypoxia-r

26	669	26.6	412	7	ADE54666 Human Pro
27	669	26.6	412	7	ADD90600 Human BHL
28	669	26.6	412	8	ADL82901 Human PRO
29	669	26.6	412	8	ADN03704 Antipsori
30	669	26.6	412	8	ADO19143 Human PRO
31	669	26.6	412	8	ADO19145 Human PRO
32	669	26.6	425	8	ADP45583 Human col
33	669	26.6	425	8	ADP45584 Human col
34	638	25.3	379	8	ADP45580 Human col
35	638	25.3	379	8	ADP45579 Human col
36	606.5	24.1	432	4	AAU30779 Novel hum
37	458.5	18.2	88	4	AAU32415 Novel hum
38	417.5	16.6	109	5	ADK34819 Novel hum
39	404	16.1	204	7	ADC3242 Human nov
40	333.5	13.2	207	4	AAU30778 Novel hum
41	269	10.7	54	5	ABP52802 Medane an
42	260	10.3	54	5	ABP52803 Murine ca
43	220.5	8.8	339	4	AAU72376 Murine ca
44	220	8.7	390	3	AAU42083 Human ORF
45	207	8.2	337	4	AAU47269 Human gri

ALIGNMENTS

RESULT 1
AAB70693
ID AAB70693 standard; protein; 484 AA.

AC AAB70693;
17-MAY-2001 (first entry)
Human DEC2b protein sequence SEQ ID NO:12.

DEC2a; DEC2b; BHLH type transcription factor; DEC2; DEC1;
basic helix loop helix protein; cell differentiation; proliferation.

OS Homo sapiens.

XX WO200114551-A1.

XX 01-MAR-2001.

XX 19-JUN-2000; 2000WO-JP003991.

XX 19-AUG-1999; 99JP-00233286.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Fujimoto K, Shin M, Kato Y;

XX WPI; 2001-202935/20.

XX N-PSDB; AAF74776.

DEC2 is a basic helix loop helix protein of the DEC family for use in development of drugs for treatment of disorders of cell differentiation and proliferation.

XX Claim 2; Page 69-72; 83pp; Japanese.

XX The present invention describes a basic helix loop helix (BHLH) type transcription factor designated DEC2. DEC2 can be used as a tool in the development of drugs for the treatment and prevention of disorders involving cell differentiation and proliferation. The present sequence represents the specifically claimed human DEC2b protein, as given in the present invention

XX Sequence 484 AA;

Query Match 100.0%; Score 2517; DB 4; Length 484;

Best Local Similarity 100.0%; Pred. No. 1.7e-187;

Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PD 19-SEP-2002.
 XX PF
 XX 14-MAR-2002; 2002WO-US007826.
 XX
 PR 14-MAR-2001; 2001US-0276025P.
 PR 14-MAR-2001; 2001US-0276026P.
 PR 10-AUG-2001; 2001US-0311732P.
 PR 19-SEP-2001; 2001US-0323580P.
 PR 26-SEP-2001; 2001US-0324967P.
 PR 26-SEP-2001; 2001US-0325102P.
 PR 26-SEP-2001; 2001US-0325149P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
 PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX
 XX WPI; 2002-723277/78.
 DR N-PSDB; ABS76400.
 DR
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient.
 XX
 PS Disclosure; Page 187-188; 481pp; English.
 XX
 CC The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. non-tuberculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining
 CC whether ovarian cancer has metastasized or is likely to metastasize,
 CC selecting a composition for inhibiting ovarian cancer, assessing the
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
 CC cancer or at risk of developing ovarian cancer. The present amino acid
 CC sequence represents one of the ovarian cancer markers described in the
 CC invention
 XX
 SQ Sequence 482 AA;
 Query Match 99.2%; Score 2498; DB 5; Length 482;
 Best Local Similarity 99.6%; Pred. No. 5 1e-186;
 Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 1 MDEGPHLQERQLLEHRDFIGLDYSLYMKPKRSMKRDYKLPRLHLEKGRD 60
 DB 1 MDEGPHLQERQLLEHRDFIGLDYSLYMKPKRSMKRDYKLPRLHLEKGRD 58
 QY 61 RINECIAQKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIIALQNGER 120
 DB 59 RINECIAQKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIIALQNGER 118
 QY 121 SLKSPISQSLDAPHSGFQCAKEVLYLSRFESWTPREPRVCQLINHLHAVATQFLPTPQ 180
 DB 119 SLKSPISQSLDAPHSGFQCAKEVLYLSRFESWTPREPRVCQLINHLHAVATQFLPTPQ 178

QY 181 LLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVFVITQRTOPSAELAENDTDT 240
 DB 179 LLTQOVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVFVITQRTOPSAELAENDTDT 238
 QY 241 DSGYGGAEAREPDREKKGAGASRVTIKQSPGCBEDSPAPKRMKLDSEGGSGGGGGGAA 300
 DB 239 DSGYGGAEAREPDREKKGAGASRVTIKQSPGCBEDSPAPKRMKLDSEGGSGGGGGGAA 298
 QY 301 AAAAALLGPDFAAAAALLRPDAALLSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPSA 360
 DB 299 AAAAALLGPDFAAAAALLRPDAALLSLVAFGGGGGAPFPQAAAAAPFCLPFCFLSPSA 358
 QY 361 AAAYVQFLDKSGLEKLYIPAAAAPEPLYPGIPAPAAAAAARAAAAAARAAAAAARAAAAA 420
 DB 359 AAAYVQFLDKSGLEKLYIPAAAAPEPLYPGIPAPAAAAAARAAAAAARAAAAAARAAAAA 418
 QY 421 LSPPEKAGAAAAATLLPHEVAPLGAHPQHPHGRTHLPFAGPREPGNPRESSAQEDPSQPG 480
 DB 419 LSPPEKAGAAAAATLLPHEVAPLGAHPQHPHGRTHLPFAGPREPGNPRESSAQEDPSQPG 478
 QY 481 KEAP 484
 DB 479 KEAP 482
 RESULT 4
 ABO14645
 ID ABO14645 standard; protein; 482 AA.
 XX
 AC ABO14645;
 XX
 DT 25-AUG-2003 (first entry)
 XX
 DE Novel human protein #18.
 XX
 KW Human; NOV; Gene therapy; endocrine related disease; diabetes;
 KW metabolism-related disease; obesity; central nervous system disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO2003023002-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 09-SEP-2002; 2002WO-US028539.
 XX
 PR 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318130P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 17-SEP-2001; 2001US-0322636P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324969P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324980P.
 PR 17-APR-2002; 2002US-0373212P.
 PR 06-SEP-2002; 2002US-00236177.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
 PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
 PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
 PI Lepley DM, Edinger SR, Burgess CE;
 XX WPI; 2003-313242/30.
 DR N-PSDB; ACD19338.

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
 PT and polynucleotides, useful in gene therapy, e.g. for treating or
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
 PT stroke or infections.

XX Claim 1; Page 125; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX
 CC polypeptide, nucleic acid and antibody are useful as therapeutics,
 CC particularly in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, which includes a pathology associated
 CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
 CC therapy for treating the disease or condition. In particular, the NOVX
 CC polypeptide or polynucleotide is useful for treating endocrine/
 CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
 CC These are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications, and for monitoring the effects of drugs during clinical
 CC trials. This is the amino acid sequence of a novel human NOV protein
 XX SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 6; Length 482;
 Best Local Similarity 99.8%; Pred. No. 5.1e-186;
 Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITKVSQTYKLPRLIEKKRD 60
 DB 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITK--DTYKLPRLIEKKRD 58
 QY 61 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLELTLTKHLKALTALTEQHQKIALLQNGER 120
 DB 59 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLELTLTKHLKALTALTEQHQKIALLQNGER 118
 QY 121 SLKSPQISDLDAFHSFGFTCAKEVQLSRFSWTPREPCVQLINHLHAVAQFLPTPQ 180
 DB 119 SLKSPQISDLDAFHSFGFTCAKEVQLSRFSWTPREPCVQLINHLHAVAQFLPTPQ 178
 QY 179 LTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAEIAAENDTDT 238
 QY 241 DSGYGGAEARPDREKKGAGASRVITKQEPGEDSPAPKMKLDSRGSGGGGGGAA 300
 DB 239 DSGYGGAEARPDREKKGAGASRVITKQEPGEDSPAPKMKLDSRGSGGGGGGAA 298
 QY 301 AAAAAALLGDPDPAALALLRPAALLSSLVAFGGGGGAPFPQAAAAAPCLPFCFLSPSA 360
 DB 299 AAAAAALLGDPDPAALALLRPAALLSSLVAFGGGGGAPFPQAAAAAPCLPFCFLSPSA 358
 QY 361 AAAYVQFDLKSLEKYLPAALAAAPFLLYPGIPAPAAAAAALAAAAAFAFLCSSV 420
 DB 359 AAAYVQFDLKSLEKYLPAALAAAPFLLYPGIPAPAAAAAALAAAAAFAFLCSSV 418
 QY 421 LSPPEKAGAAAAATLLPHEVAPLIGAPHQHPHGRTHLPFAGPREPCNPRESSAQEDPSQPG 480
 DB 419 LSPPEKAGAAAAATLLPHEVAPLIGAPHQHPHGRTHLPFAGPREPCNPRESSAQEDPSQPG 478

QY 481 KEAP 484
 DB 479 KEAP 482

RESULT 5
 ABR58470
 ID ABR58470 standard; protein; 482 AA.
 AC ABR58470;
 XX 07-JUL-2003 (first entry)
 DE Human bHLH protein DEC2 from ovarian cancer clone O1668P.
 KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.
 OS Homo sapiens.
 PN WO2003029468-A1.
 PD 10-APR-2003.
 PF 02-OCT-2002; 2002WO-US031467.
 PR 02-OCT-2001; 2001US-0327135P.
 PR 30-MAY-2002; 2002US-0384531P.
 PA (CORI-) CORIXA CORP.
 PI Algate PA, Mannion J;
 XX WPI; 2003-372001/35.

XX New polynucleotide and polypeptide useful for diagnosing and/or treating
 PT cancer, particularly ovarian cancer, and as a vaccine.
 XX Claim 2; Page 156-157; 169pp; English.
 CC The invention relates to a novel isolated polynucleotide. The
 CC polynucleotides of the invention have cytostatic activity, and may have a
 CC use in gene therapy, and in a vaccine. The composition and methods are
 CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
 CC The composition may also be used as a vaccine to prevent cancer. The
 CC present sequence is used in the exemplification of the invention
 XX SQ Sequence 482 AA;

Query Match 99.2%; Score 2498; DB 6; Length 482;
 Best Local Similarity 99.8%; Pred. No. 5.1e-186;
 Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITKVSQTYKLPRLIEKKRD 60
 DB 1 MDEGIPHLQERQLLEHRDFIGLDYSSLYMCKPKRSMKRDITK--DTYKLPRLIEKKRD 58
 QY 61 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLELTLTKHLKALTALTEQHQKIALLQNGER 120
 DB 59 RINECIAQLKDLLPEHLKLTTLTGHLKAVVLELTLTKHLKALTALTEQHQKIALLQNGER 118
 QY 121 SLKSPQISDLDAFHSFGFTCAKEVQLSRFSWTPREPCVQLINHLHAVAQFLPTPQ 180
 DB 119 SLKSPQISDLDAFHSFGFTCAKEVQLSRFSWTPREPCVQLINHLHAVAQFLPTPQ 178
 QY 181 LTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAEIAAENDTDT 240
 DB 179 LTQQVPLSKGTGAPSAAGSAAAPCLERAGQKLEPLAYCVPIQRTQPSAEIAAENDTDT 238
 QY 241 DSGYGGAEARPDREKKGAGASRVITKQEPGEDSPAPKMKLDSRGSGGGGGGAA 300
 DB 239 DSGYGGAEARPDREKKGAGASRVITKQEPGEDSPAPKMKLDSRGSGGGGGGAA 298

Best Local Similarity 99.6%; Pred. No. 5.1e-186;		Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;	
QY	1	MDEGPHLQRLLEHRDFIGLDYSSLYMCKPKRSMKRD	TKVSDTYKLPRLIEKKRRD 60
Db	1	MDEGPHLQRLLEHRDFIGLDYSSLYMCKPKRSMKRD	DTK--DTYKLPRLIEKKRRD 58
QY	61	RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT	LKHLKALTALTEQOQHKIATLONGER 120
Db	59	RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT	LKHLKALTALTEQOQHKIATLONGER 118
QY	121	SLKSPIQSDLDAPHSGFQTC	AKVLYLSRFSWTPREPCVQLINHLHAVA
Db	119	SLKSPIQSDLDAPHSGFQTC	AKVLYLSRFSWTPREPCVQLINHLHAVA
QY	181	LLTQOVPLSKGTGAPSAAGSAAPCLERAGQKLE	PLAYCVPVITQTPSAELAAENDTDT 240
Db	179	LLTQOVPLSKGTGAPSAAGSAAPCLERAGQKLE	PLAYCVPVITQTPSAELAAENDTDT 238
QY	241	DSGYGGEAEARPDREKKGAGASRVTIKQPPGED	SPAPKRMKLD
Db	239	DSGYGGEAEARPDREKKGAGASRVTIKQPPGED	SPAPKRMKLD
QY	301	AAAAALLGPPDAAAAALLRPDAALLSLVAF	GGGGAPFPQAAAAAPCLPFCFLSPSA 360
Db	299	AAAAALLGPPDAAAAALLRPDAALLSLVAF	GGGGAPFPQAAAAAPCLPFCFLSPSA 358
QY	361	AAAYVOPFLDKSLEKLYPAAAAAPFLLY	PGIPAPAAAAA
Db	359	AAAYVOPFLDKSLEKLYPAAAAAPFLLY	PGIPAPAAAAA
QY	421	LSPPEKAGAAATLFPHEVAPLGP	HPHQHRTLFPAGPREPCNPSSAQEDPSQPG 480
Db	419	LSPPEKAGAAATLFPHEVAPLGP	HPHQHRTLFPAGPREPCNPSSAQEDPSQPG 478
QY	481	KEAP 484	
Db	479	KEAP 482	
RESULT 9			
ABJ04644			
ID	ABJ04644	standard; protein; 767 AA.	
AC	AC		
XX	ABJ04644;		
DT	DT		
XX	11-OCT-2002 (first entry)		
DE	Protein of NOVX 5 SEQ ID No 10.		
XX			
KW	Cytostatic; antidiabetic; anorectic; metabolic; neurotic; norepinephrine; antiparkinsonian; anticonvulsant; cerebroprotective; neuroprotective; tranquilizer; neuroleptic; antidiabetic; antitumor; anti-HIV; antiallergic; antirheumatic; antidiabetic; NOVX; diabetes; metabolic disorder; obesity; infectious disease; Alzheimer's disease; anorexia; neurodegenerative disorder; Parkinson's disease; obesity; immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease; metabolic syndrome X; wasting disorder; cancer; neurological disorder; epilepsy; stroke; mental disorder; schizophrenia; schizophrenia; goiter; vesicular transport; cystic fibrosis; gastrointestinal disorder; diabetes mellitus; ulcerative colitis; AIDS; allergic reaction; multiple sclerosis; rheumatoid arthritis; transgenic animal; gene therapy.		
OS	Unidentified.		
XX			
PN	WO200246409-A2.		
XX			
PD	13-JUN-2002.		
XX			
PF	06-DEC-2001; 2001WO-US046586.		
XX			
PR	06-DEC-2000; 2000US-0251660P.		

PR	12-DEC-2000; 2000US-0255032P.		
PR	08-JAN-2001; 2001US-0260326P.		
PR	24-JAN-2001; 2001US-0263800P.		
PR	20-FEB-2001; 2001US-0269942P.		
PR	24-APR-2001; 2001US-0286183P.		
PR	20-AUG-2001; 2001US-0313627P.		
PR	12-SEP-2001; 2001US-0318712P.		
XX			
PA	(CURA-) CURAGEN CORP.		
XX			
PI	Guo X, Li L, Patturajan M, Shinkets RA, Casman SU, Malyankar UM;		
PI	Tchernev VT, Vernet CM, Spytek KA, Shenoy SG, Alsobrook JP;		
PI	Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;		
PI	Colman SD, Bisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;		
XX			
DR	WPI; 2002-547774/58.		
DR	N-PSDB; ABT05457.		
XX			
PT	Novel isolated polypeptide, designated NOVX, useful for treating or		
PT	preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and		
PT	metabolic, neurodegenerative, immune and hematopoietic disorders.		
XX			
PS	Claim 1; Page 44; 421pp; English.		
XX			
CC	The invention relates to an isolated polypeptide, designated NOVX,		
CC	comprising a sequence fully defined in the specification. The isolated		
CC	protein, its encoding polynucleotide or an antibody created from the		
CC	protein is useful in the manufacture of a medicament for treating a		
CC	syndrome associated with a human disease, preferably a NOVX-associated		
CC	disorder, or for treating or preventing a NOVX-associated disorder in a		
CC	subject, preferably human. The isolated protein, its encoding		
CC	polynucleotide or an antibody created from the protein are also useful		
CC	for treating or preventing metabolic disorders, diabetes, obesity,		
CC	infectious disease, anorexia, neurodegenerative disorder, Alzheimer's		
CC	disease, Parkinson's disorder, immune disorders, haematopoietic		
CC	disorders, and various dyslipidaemias, metabolic disturbances associated		
CC	with obesity, the metabolic syndrome X, wasting disorders associated with		
CC	chronic diseases, and cancer. The isolated protein, its encoding		
CC	polynucleotide or an antibody created from the protein are useful for		
CC	treating or preventing neurological disorders such as epilepsy, stroke,		
CC	mental disorders including mood, anxiety, schizophrenic disorders,		
CC	disorders of vesicular transport such as cystic fibrosis, diabetes,		
CC	melittus, goiter, gastrointestinal disorders including ulcerative		
CC	colitis, other conditions associated with abnormal vesicle trafficking		
CC	including AIDS, allergic reactions, multiple sclerosis and rheumatoid		
CC	arthritis. A cell comprising the vector of the invention is useful for		
CC	producing non-human transgenic animals. The polynucleotide of the		
CC	invention can be used to treat disorders by gene therapy. This sequence		
CC	represents one of the isolated NOVX proteins of the invention		
XX			
SQ	Sequence 767 AA;		
Query Match 99.2%; Score 2498; DB 5; Length 767;		Best Local Similarity 99.6%; Pred. No. 9.3e-186;	
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 1;			
QY	1	MDEGPHLQRLLEHRDFIGLDYSSLYMCKPKRSMKRD	TKVSDTYKLPRLIEKKRRD 60
Db	286	MDEGPHLQRLLEHRDFIGLDYSSLYMCKPKRSMKRD	DTK--DTYKLPRLIEKKRRD 343
QY	61	RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT	LKHLKALTALTEQOQHKIATLONGER 120
Db	344	RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELT	LKHLKALTALTEQOQHKIATLONGER 403
QY	121	SLKSPIQSDLDAPHSGFQTC	AKVLYLSRFSWTPREPCVQLINHLHAVA
Db	404	SLKSPIQSDLDAPHSGFQTC	AKVLYLSRFSWTPREPCVQLINHLHAVA
QY	181	LLTQOVPLSKGTGAPSAAGSAAPCLERAGQKLE	PLAYCVPVITQTPSAELAAENDTDT 240
Db	464	LLTQOVPLSKGTGAPSAAGSAAPCLERAGQKLE	PLAYCVPVITQTPSAELAAENDTDT 523
QY	241	DSGYGGEAEARPDREKKGAGASRVTIKQPPGED	SPAPKRMKLD

Db 524 DSGYGEAEARPDREKKGAGASRVTIKQEPGDSAPAKRMKLDLSRGSGGGGGGNA 583
QY 301 AAAAAALLGPDPAALALRPDAALLSSLVAFGGGGGAPPQPAALAPCLPFCFLSPSA 360
Db 584 AAAAAALLGPDPAALALRPDAALLSSLVAFGGGGGAPPQPAALAPCLPFCFLSPSA 643
QY 361 AAVVOPFLDKSGLEKYLPAALAPFLLYPCIPAPAAAAAALAAAAAALPCLSSV 420
Db 644 AAVVOPFLDKSGLEKYLPAALAPFLLYPCIPAPAAAAAALAAAAAALPCLSSV 703
QY 421 LSPPEKAGAAAATLLPHEVAPLGAHPHQPHTHLPAGPREGNPNPSSAQEDPSQPG 480
Db 704 LSPPEKAGAAAATLLPHEVAPLGAHPHQPHTHLPAGPREGNPNPSSAQEDPSQPG 763
QY 481 KEAP 484
Db 764 KEAP 767

RESULT 10
AAB70694
ID AAB70694 standard; protein; 410 AA.
XX AAB70694;
AC AC
XX AC
XX AC
DT 17-MAY-2001 (first entry)
DE Mouse DEC2a protein sequence SEQ ID NO:14.
XX DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation.
XX Mus musculus.
XX WO200114551-A1.
XX 01-MAR-2001.
XX 19-JUN-2000; 2000WO-JP003991.
XX 19-AUG-1999; 99JP-00233286.
XX (CHUS) CHUGAI SEIYAKU KK.
XX Fujimoto K, Shin M, Kato Y;
XX WPI; 2001-202935/20.
XX N-PSDB; AAF74777.
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
and proliferation.
XX Claim 2; Page 77-80; 83pp; Japanese.
XX The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC represents the specifically claimed mouse DEC2a protein, as given in the
CC present invention
XX Sequence 410 AA;
SQ

Query Match 54.9%; Score 1634; DB 4; Length 410;
Best Local Similarity 72.0%; Pred. No. 9.3e-119;
Matches 350; Conservative 21; Mismatches 37; Indels 78; Gaps 16;
QY 1 MDEGIPHLQERQLLEHRRDFGLDYSSLYMCKPKRSMKRDYKVDYKLPRLIEKKRD 60
Db 1 MDEGIPHLQERQLLEHRRDFGLDYSSLYMCKPKRSLKRDYK--DYKLPRLIEKKRD 58

QY 61 RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELTTLKHLKALTALTEQHQKIALLQNGER 120
Db 59 RINECIAQLKDLLPEHLKLTTLGLHLEKAVVLELTTLKHLKALTALTEQHQKIALLQNGER 118
QY 121 SLKSPIQSDLDADHSGFQTCACEVLYLSRPFESWTREPRCQVLIHNLHNAVATQFLPTQ 180
Db 119 SLKSPVQADLDADHSGFQTCACEVLYLSRPFESWTREPRCQVLSHLHNAVAT-----Q 172
QY 181 LLTQCVPLSKGTG-APSAAGSAAAPCLERAGOKLEPLAYCVPIQRTQPSAEALAAENDTD 239
Db 173 LLTPQVSGRGSGRAPCSAGAAAASGPER-----VAVCVPIQRTQPGTE--PEHDTD 223
QY 240 TDSYGYGEAEARPDREKKGAGASRVTIKQEPGDSAPAKRMKLDLSRGSGGGGGGA 299
Db 224 TDSYGYGEAE-----QGRAAVKQEPGDSAPAKRMKLEARG----- 260
QY 300 AAAAAALLGPDPAALALRPDAALLSSLVAFGGGGGAPPQPAALAPCLPFCFLSPS 359
Db 261 -----ALLGPEP-----ALLGSLVAL--GGAPPAQP--AAAPCLPFCFLSPS 300
QY 360 AAVVOPFLDKSGLEKYLPAALAPFLLYPCIPAPAAAAAALAAAAAALPCLSS 419
Db 301 -AAAVVOPWLDKSLDKLYLP-AAAAFFLLYPCIP-----AAAAAALPCLSS 350
QY 420 VLSPPPEKAGA-AAATLLPHEVAPLGAHPHQPHTHLPAGPREGNPNPSSAQEDPSQ 478
Db 351 VLSPPPEKAGATAGAPFLAHEVAPPGPLRQPHAHSRHTLFR-----VNPES- QEDATQ 404
QY 479 PGKEAP 484
Db 405 PAKDAP 410

RESULT 11
ADE54660
ID ADE54660 standard; protein; 410 AA.
XX ADE54660;
AC ADE54660;
XX 29-JAN-2004 (first entry)
XX Rat Protein O35779, SEQ ID NO 465.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENEBANK; O35779.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI), and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 410 AA;

Query Match 64.5%; Score 1624; DB 7; Length 410;
Best Local Similarity 71.6%; Pred. No. 5.6e-118;
Matches 348; Conservative 23; Mismatches 37; Indels 78; Gaps 16;
QY 1 MDEGPHLQEROLLEHRDFTGLDYSSLYMCKPKSRKRDITKVDYKLPRLIEKKRD 60
DB 1 MDEGPHLQEROLLEHRDFTGLDYSSLYMCKPKSRKRDITK--DYKLPRLIEKKRD 58
QY 61 RINECIAQLKDLPEHLKUTTLGHLKAVVLETLKHLKALTALTEQOQKIIAQLNGER 120
DB 59 RINECIAQLKDLPEHLKUTTLGHLKAVVLETLKHLKALTALTEQOQKIIAQLNGER 118
QY 121 SLKSPQSDLDAPHSGFQTCARVLYLSRFESWTPREPRCVQLINHLHAVATQFLPTQ 180
DB 119 SLKSPVQADLDAPHSGFQTCARVLYLSRFESWTPREPRCAQLVSHLHAVAT-----Q 172
QY 181 LLTQVPLSKGTG-APSAAGSAAAPCLERAGOKLELAYCVPIQRTQPSAELAEENDTD 239
DB 173 LLTQVTPGPGRAPCSAGAAASGSR-----VARCVPIQRTQGTB--PEHDTD 223
QY 240 TDSGYGGEAEAPRPREKKGAGASRVTIKQEPFGEDSPAPKMKLDSRGSGGGGPGGA 299
DB 224 TDSGYGGEAE-----QGRAAVKQEPFGDPSAAPKRLKLEARG----- 260
QY 300 AAAAALLGPDPAALRPPDALLSLVAGGGGAPFPQPAALAAFPCLPFLSPS 359
DB 261 -----ALLGEP-----ALLGSLVAL--GGGAPFAQP--AAAPFCLPFLVLLSPS 300
QY 360 AAAAYVQPFDLKSGLEKLYLPAAAPFPFLVPGIPAPAAAAAALAAAAAAPPCLSS 419
DB 301 -AAAYVQPLDKSLDKLYLP-AAAPFLLYFIP-----AAAAAALAAFPCLSS 350
QY 420 VLSFPPEKAGAAA-ATLLPHEVAPLGAHPQHPGHGTHLPFGAPRPPGNPSSAGDEPSQ 478
DB 351 VLSFPPEKAGSAGAPFAHEVAPPGSLRFPQHAHSRTHLPHA-----VNPSS-QEDATQ 404
QY 479 PGKEAP 484
DB 405 PAKDAP 410

RESULT 12

AB014646

ID AB014646 standard; protein; 203 AA.

XX

AC
XX
DT
XX
DE
XX
KW
KW
KW
KW
KW
XX
OS
PN
PD
PF
XX
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PR
PA
XX
PI
PI
PI
PI
PI
XX
DR
DR
XX
PT
PT
PT
PT
XX
XX
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC

AB014646;
25-AUG-2003 (first entry)
Novel human protein #19.
Human; NOV; gene therapy; endocrine related disease; diabetes;
metabolism-related disease; obesity; central nervous system disorder;
Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
schizophrenia; depression; autoimmune disorder; inflammatory disorder;
psoriasis; allergy; lupus erythematosus; asthma; cancer;
inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
lung disease; emphysema; obstructive pulmonary disease; haemophilia;
stroke; infection.
Homo sapiens.
WO2003023002-A2.
20-MAR-2003.
09-SEP-2002; 2002WO-US028539.
07-SEP-2001; 2001US-0318120P.
07-SEP-2001; 2001US-0318130P.
10-SEP-2001; 2001US-0318430P.
17-SEP-2001; 2001US-0322636P.
17-SEP-2001; 2001US-0322781P.
17-SEP-2001; 2001US-0322816P.
17-SEP-2001; 2001US-0322817P.
19-SEP-2001; 2001US-0323519P.
20-SEP-2001; 2001US-0323631P.
20-SEP-2001; 2001US-0323636P.
25-SEP-2001; 2001US-0324969P.
25-SEP-2001; 2001US-0325091P.
26-SEP-2001; 2001US-0324930P.
17-APR-2002; 2002US-0373212P.
06-SEP-2002; 2002US-00236177.
(CURA-) CURAGEN CORP.
Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
Gerlach VL, Vernet CM, Berghs C, Rothenberg ME, Guo X;
Shankets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
Lepley DM, Edinger SR, Burgess CE;
WPI; 2003-313242/30.
N-PSDB; ACD19339.
New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
and polynucleotides, useful in gene therapy, e.g. for treating or
preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
stroke or infections.
Claim 1; Page 126; 586pp; English.
The invention describes a new isolated polypeptide (NOVX). The NOVX
polypeptide, nucleic acid and antibody are useful as therapeutics,
particularly in the manufacture of a medicament for treating a syndrome
associated with a human disease, which includes a pathology associated
with NOVX polypeptide. The DNA encoding the protein is useful in gene
therapy for treating the disease or condition. In particular, the NOVX
polypeptide or polynucleotide is useful for treating endocrine/
metabolism-related diseases (e.g. obesity or diabetes), central nervous
system disorders (e.g. Alzheimer's disease, Parkinson's disease,
epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
asthma, inflammatory bowel disease, rheumatoid arthritis or
osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
prostate or brain cancers, or melanoma), liver diseases (e.g. liver

CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
 CC These are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications, and for monitoring the effects of drugs during clinical
 CC trials. This is the amino acid sequence of a novel human NOV protein
 XX
 SQ Sequence 203 AA;

Query Match 36.0%; Score 905; DB 6; Length 203;
 Best Local Similarity 88.2%; Pred. No. 2.5e-62;
 Matches 179; Conservative 2; Mismatches 12; Indels 10; Gaps 2;
 QY 1 MDEGIPHLERQLLEHRDFIGLDYSSLYCKPKRSMKRDYKVDYKLPKHLIEKKRD 60
 DB 3 MDEGIPHLERQLPEHRDFIGLDYSSLYCKPKRSMKRDYK--DTYKLPKHLIEKKRD 60
 QY 61 RINECTAQLKDLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQHQKIIALQNGER 120
 DB 61 RINECTAQLKDLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQHQKIIALQNGER 120
 QY 121 SLKSPISQSLDADFSGFQTCACEVLYLSRFSWTPREPRCVQLINHLHVAATQFLP--- 177
 DB 121 SLKSPISQSLDADFSGFQTCACEVLYLSRFSWTPREPRCVQLINHLHVAATQFLPAG 180
 QY 178 -----TPQLLTQVPLSKGTGAP 195
 DB 181 PREPGNPSSAQEDPSQPGKEAP 203

RESULT 13
 AAU16188
 ID AAU16188 standard; protein; 165 AA.
 XX AC AAU16188;
 XX DT 07-NOV-2001 (first entry)
 XX DE Human novel secreted protein, Seq ID 1141.
 XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
 KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;
 KW preservative; antiproliferative.
 XX OS Homo sapiens.
 XX PN W0200155322-A2.
 XX PD 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US001341.
 XX PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226682P.
 PR 22-AUG-2000; 2000US-0226683P.
 PR 22-AUG-2000; 2000US-0226684P.
 PR 22-AUG-2000; 2000US-0226685P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0226687P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0226689P.
 PR 22-AUG-2000; 2000US-0226690P.
 PR 22-AUG-2000; 2000US-0226691P.
 PR 22-AUG-2000; 2000US-0226692P.
 PR 22-AUG-2000; 2000US-0226693P.
 PR 22-AUG-2000; 2000US-0226694P.
 PR 22-AUG-2000; 2000US-0226695P.
 PR 22-AUG-2000; 2000US-0226696P.
 PR 22-AUG-2000; 2000US-0226697P.
 PR 22-AUG-2000; 2000US-0226698P.
 PR 22-AUG-2000; 2000US-0226699P.
 PR 22-AUG-2000; 2000US-0226700P.
 PR 22-AUG-2000; 2000US-0226701P.
 PR 22-AUG-2000; 2000US-0226702P.
 PR 22-AUG-2000; 2000US-0226703P.
 PR 22-AUG-2000; 2000US-0226704P.
 PR 22-AUG-2000; 2000US-0226705P.
 PR 22-AUG-2000; 2000US-0226706P.
 PR 22-AUG-2000; 2000US-0226707P.
 PR 22-AUG-2000; 2000US-0226708P.
 PR 22-AUG-2000; 2000US-0226709P.
 PR 22-AUG-2000; 2000US-0226710P.
 PR 22-AUG-2000; 2000US-0226711P.
 PR 22-AUG-2000; 2000US-0226712P.
 PR 22-AUG-2000; 2000US-0226713P.
 PR 22-AUG-2000; 2000US-0226714P.
 PR 22-AUG-2000; 2000US-0226715P.
 PR 22-AUG-2000; 2000US-0226716P.
 PR 22-AUG-2000; 2000US-0226717P.
 PR 22-AUG-2000; 2000US-0226718P.
 PR 22-AUG-2000; 2000US-0226719P.
 PR 22-AUG-2000; 2000US-0226720P.
 PR 22-AUG-2000; 2000US-0226721P.
 PR 22-AUG-2000; 2000US-0226722P.
 PR 22-AUG-2000; 2000US-0226723P.
 PR 22-AUG-2000; 2000US-0226724P.
 PR 22-AUG-2000; 2000US-0226725P.
 PR 22-AUG-2000; 2000US-0226726P.
 PR 22-AUG-2000; 2000US-0226727P.
 PR 22-AUG-2000; 2000US-0226728P.
 PR 22-AUG-2000; 2000US-0226729P.
 PR 22-AUG-2000; 2000US-0226730P.
 PR 22-AUG-2000; 2000US-0226731P.
 PR 22-AUG-2000; 2000US-0226732P.
 PR 22-AUG-2000; 2000US-0226733P.
 PR 22-AUG-2000; 2000US-0226734P.
 PR 22-AUG-2000; 2000US-0226735P.
 PR 22-AUG-2000; 2000US-0226736P.
 PR 22-AUG-2000; 2000US-0226737P.
 PR 22-AUG-2000; 2000US-0226738P.
 PR 22-AUG-2000; 2000US-0226739P.
 PR 22-AUG-2000; 2000US-0226740P.
 PR 22-AUG-2000; 2000US-0226741P.
 PR 22-AUG-2000; 2000US-0226742P.
 PR 22-AUG-2000; 2000US-0226743P.
 PR 22-AUG-2000; 2000US-0226744P.
 PR 22-AUG-2000; 2000US-0226745P.
 PR 22-AUG-2000; 2000US-0226746P.
 PR 22-AUG-2000; 2000US-0226747P.
 PR 22-AUG-2000; 2000US-0226748P.
 PR 22-AUG-2000; 2000US-0226749P.
 PR 22-AUG-2000; 2000US-0226750P.
 PR 22-AUG-2000; 2000US-0226751P.
 PR 22-AUG-2000; 2000US-0226752P.
 PR 22-AUG-2000; 2000US-0226753P.
 PR 22-AUG-2000; 2000US-0226754P.
 PR 22-AUG-2000; 2000US-0226755P.
 PR 22-AUG-2000; 2000US-0226756P.
 PR 22-AUG-2000; 2000US-0226757P.
 PR 22-AUG-2000; 2000US-0226758P.
 PR 22-AUG-2000; 2000US-0226759P.
 PR 22-AUG-2000; 2000US-0226760P.
 PR 22-AUG-2000; 2000US-0226761P.
 PR 22-AUG-2000; 2000US-0226762P.
 PR 22-AUG-2000; 2000US-0226763P.
 PR 22-AUG-2000; 2000US-0226764P.
 PR 22-AUG-2000; 2000US-0226765P.
 PR 22-AUG-2000; 2000US-0226766P.
 PR 22-AUG-2000; 2000US-0226767P.
 PR 22-AUG-2000; 2000US-0226768P.
 PR 22-AUG-2000; 2000US-0226769P.
 PR 22-AUG-2000; 2000US-0226770P.
 PR 22-AUG-2000; 2000US-0226771P.
 PR 22-AUG-2000; 2000US-0226772P.
 PR 22-AUG-2000; 2000US-0226773P.
 PR 22-AUG-2000; 2000US-0226774P.
 PR 22-AUG-2000; 2000US-0226775P.
 PR 22-AUG-2000; 2000US-0226776P.
 PR 22-AUG-2000; 2000US-0226777P.
 PR 22-AUG-2000; 2000US-0226778P.
 PR 22-AUG-2000; 2000US-0226779P.
 PR 22-AUG-2000; 2000US-0226780P.
 PR 22-AUG-2000; 2000US-0226781P.
 PR 22-AUG-2000; 2000US-0226782P.
 PR 22-AUG-2000; 2000US-0226783P.
 PR 22-AUG-2000; 2000US-0226784P.
 PR 22-AUG-2000; 2000US-0226785P.
 PR 22-AUG-2000; 2000US-0226786P.
 PR 22-AUG-2000; 2000US-0226787P.
 PR 22-AUG-2000; 2000US-0226788P.
 PR 22-AUG-2000; 2000US-0226789P.
 PR 22-AUG-2000; 2000US-0226790P.
 PR 22-AUG-2000; 2000US-0226791P.
 PR 22-AUG-2000; 2000US-0226792P.
 PR 22-AUG-2000; 2000US-0226793P.
 PR 22-AUG-2000; 2000US-0226794P.
 PR 22-AUG-2000; 2000US-0226795P.
 PR 22-AUG-2000; 2000US-0226796P.
 PR 22-AUG-2000; 2000US-0226797P.
 PR 22-AUG-2000; 2000US-0226798P.
 PR 22-AUG-2000; 2000US-0226799P.
 PR 22-AUG-2000; 2000US-0226800P.
 PR 22-AUG-2000; 2000US-0226801P.
 PR 22-AUG-2000; 2000US-0226802P.
 PR 22-AUG-2000; 2000US-0226803P.
 PR 22-AUG-2000; 2000US-0226804P.
 PR 22-AUG-2000; 2000US-0226805P.
 PR 22-AUG-2000; 2000US-0226806P.
 PR 22-AUG-2000; 2000US-0226807P.
 PR 22-AUG-2000; 2000US-0226808P.
 PR 22-AUG-2000; 2000US-0226809P.
 PR 22-AUG-2000; 2000US-0226810P.
 PR 22-AUG-2000; 2000US-0226811P.
 PR 22-AUG-2000; 2000US-0226812P.
 PR 22-AUG-2000; 2000US-0226813P.
 PR 22-AUG-2000; 2000US-0226814P.
 PR 22-AUG-2000; 2000US-0226815P.
 PR 22-AUG-2000; 2000US-0226816P.
 PR 22-AUG-2000; 2000US-0226817P.
 PR 22-AUG-2000; 2000US-0226818P.
 PR 22-AUG-2000; 2000US-0226819P.
 PR 22-AUG-2000; 2000US-0226820P.
 PR 22-AUG-2000; 2000US-0226821P.
 PR 22-AUG-2000; 2000US-0226822P.
 PR 22-AUG-2000; 2000US-0226823P.
 PR 22-AUG-2000; 2000US-0226824P.
 PR 22-AUG-2000; 2000US-0226825P.
 PR 22-AUG-2000; 2000US-0226826P.
 PR 22-AUG-2000; 2000US-0226827P.
 PR 22-AUG-2000; 2000US-0226828P.
 PR 22-AUG-2000; 2000US-0226829P.
 PR 22-AUG-2000; 2000US-0226830P.
 PR 22-AUG-2000; 2000US-0226831P.
 PR 22-AUG-2000; 2000US-0226832P.
 PR 22-AUG-2000; 2000US-0226833P.
 PR 22-AUG-2000; 2000US-0226834P.
 PR 22-AUG-2000; 2000US-0226835P.
 PR 22-AUG-2000; 2000US-0226836P.
 PR 22-AUG-2000; 2000US-0226837P.
 PR 22-AUG-2000; 2000US-0226838P.
 PR 22-AUG-2000; 2000US-0226839P.
 PR 22-AUG-2000; 2000US-0226840P.
 PR 22-AUG-2000; 2000US-0226841P.
 PR 22-AUG-2000; 2000US-0226842P.
 PR 22-AUG-2000; 2000US-0226843P.
 PR 22-AUG-2000; 2000US-0226844P.
 PR 22-AUG-2000; 2000US-0226845P.
 PR 22-AUG-2000; 2000US-0226846P.
 PR 22-AUG-2000; 2000US-0226847P.
 PR 22-AUG-2000; 2000US-0226848P.
 PR 22-AUG-2000; 2000US-0226849P.
 PR 22-AUG-2000; 2000US-0226850P.
 PR 22-AUG-2000; 2000US-0226851P.
 PR 22-AUG-2000; 2000US-0226852P.
 PR 22-AUG-2000; 2000US-0226853P.
 PR 22-AUG-2000; 2000US-0226854P.
 PR 22-AUG-2000; 2000US-0226855P.
 PR 22-AUG-2000; 2000US-0226856P.
 PR 22-AUG-2000; 2000US-0226857P.
 PR 22-AUG-2000; 2000US-0226858P.
 PR 22-AUG-2000; 2000US-0226859P.
 PR 22-AUG-2000; 2000US-0226860P.
 PR 22-AUG-2000; 2000US-0226861P.
 PR 22-AUG-2000; 2000US-0226862P.
 PR 22-AUG-2000; 2000US-0226863P.
 PR 22-AUG-2000; 2000US-0226864P.
 PR 22-AUG-2000; 2000US-0226865P.
 PR 22-AUG-2000; 2000US-0226866P.
 PR 22-AUG-2000; 2000US-0226867P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0226869P.
 PR 22-AUG-2000; 2000US-0226870P.
 PR 22-AUG-2000; 2000US-0226871P.
 PR 22-AUG-2000; 2000US-0226872P.
 PR 22-AUG-2000; 2000US-0226873P.
 PR 22-AUG-2000; 2000US-0226874P.
 PR 22-AUG-2000; 2000US-0226875P.
 PR 22-AUG-2000; 2000US-0226876P.
 PR 22-AUG-2000; 2000US-0226877P.
 PR 22-AUG-2000; 2000US-0226878P.
 PR 22-AUG-2000; 2000US-0226879P.
 PR 22-AUG-2000; 2000US-0226880P.
 PR 22-AUG-2000; 2000US-0226881P.
 PR 22-AUG-2000; 2000US-0226882P.
 PR 22-AUG-2000; 2000US-0226883P.
 PR 22-AUG-2000; 2000US-0226884P.
 PR 22-AUG-2000; 2000US-0226885P.
 PR 22-AUG-2000; 2000US-0226886P.
 PR 22-AUG-2000; 2000US-0226887P.
 PR 22-AUG-2000; 2000US-0226888P.
 PR 22-AUG-2000; 2000US-0226889P.
 PR 22-AUG-2000; 2000US-0226890P.
 PR 22-AUG-2000; 2000US-0226891P.
 PR 22-AUG-2000; 2000US-0226892P.
 PR 22-AUG-2000; 2000US-0226893P.
 PR 22-AUG-2000; 2000US-0226894P.
 PR 22-AUG-2000; 2000US-0226895P.
 PR 22-AUG-2000; 2000US-0226896P.
 PR 22-AUG-2000; 2000US-0226897P.
 PR 22-AUG-2000; 2000US-0226898P.
 PR 22-AUG-2000; 2000US-0226899P.
 PR 22-AUG-2000; 2000US-0226900P.
 PR 22-AUG-2000; 2000US-0226901P.
 PR 22-AUG-2000; 2000US-0226902P.
 PR 22-AUG-2000; 2000US-0226903P.
 PR 22-AUG-2000; 2000US-0226904P.
 PR 22-AUG-2000; 2000US-0226905P.
 PR 22-AUG-2000; 2000US-0226906P.
 PR 22-AUG-2000; 2000US-0226907P.
 PR 22-AUG-2000; 2000US-0226908P.
 PR 22-AUG-2000; 2000US-0226909P.
 PR 22-AUG-2000; 2000US-0226910P.
 PR 22-AUG-2000; 2000US-0226911P.
 PR 22-AUG-2000; 2000US-0226912P.
 PR 22-AUG-2000; 2000US-0226913P.
 PR 22-AUG-2000; 2000US-0226914P.
 PR 22-AUG-2000; 2000US-0226915P.
 PR 22-AUG-2000; 2000US-0226916P.
 PR 22-AUG-2000; 2000US-0226917P.
 PR 22-AUG-2000; 2000US-0226918P.
 PR 22-AUG-2000; 2000US-0226919P.
 PR 22-AUG-2000; 2000US-0226920P.
 PR 22-AUG-2000; 2000US-0226921P.
 PR 22-AUG-2000; 2000US-0226922P.
 PR 22-AUG-2000; 2000US-0226923P.
 PR 22-AUG-2000; 2000US-0226924P.
 PR 22-AUG-2000; 2000US-0226925P.
 PR 22-AUG-2000; 2000US-0226926P.
 PR 22-AUG-2000; 2000US-0226927P.
 PR 22-AUG-2000; 2000US-0226928P.
 PR 22-AUG-2000; 2000US-0226929P.
 PR 22-AUG-2000; 2000US-0226930P.
 PR 22-AUG-2000; 2000US-0226931P.
 PR 22-AUG-2000; 2000US-0226932P.
 PR 22-AUG-2000; 2000US-0226933P.
 PR 22-AUG-2000; 2000US-0226934P.
 PR 22-AUG-2000; 2000US-0226935P.
 PR 22-AUG-2000; 2000US-0226936P.
 PR 22-AUG-2000; 2000US-0226937P.
 PR 22-AUG-2000; 2000US-0226938P.
 PR 22-AUG-2000; 2000US-0226939P.
 PR 22-AUG-2000; 2000US-0226940P.
 PR 22-AUG-2000; 2000US-0226941P.
 PR 22-AUG-2000; 2000US-0226942P.
 PR 22-AUG-2000; 2000US-0226943P.
 PR 22-AUG-2000; 2000US-0226944P.
 PR 22-AUG-2000; 2000US-0226945P.
 PR 22-AUG-2000; 2000US-0226946P.
 PR 22-AUG-2000; 2000US-0226947P.
 PR 22-AUG-2000; 2000US-0226948P.
 PR 22-AUG-2000; 2000US-0226949P.
 PR 22-AUG-2000; 2000US-0226950P.
 PR 22-AUG-2000; 2000US-0226951P.
 PR 22-AUG-2000; 2000US-0226952P.
 PR 22-AUG-2000; 2000US-0226953P.
 PR 22-AUG-2000; 2000US-0226954P.
 PR 22-AUG-2000; 2000US-0226955P.
 PR 22-AUG-2000; 2000US-0226956P.
 PR 22-AUG-2000; 2000US-0226957P.
 PR 22-AUG-2000; 2000US-0226958P.
 PR 22-AUG-2000; 2000US-0226959P.
 PR 22-AUG-2000; 2000US-0226960P.
 PR 22-AUG-2000; 2000US-0226961P.
 PR 22-AUG-2000; 2000US-0226962P.
 PR 22-AUG-2000; 2000US-0226963P.
 PR 22-AUG-2000; 2000US-0226964P.
 PR 22-AUG-2000; 2000US-0226965P.
 PR 22-AUG-2000; 2000US-0226966P.
 PR 22-AUG-2000; 2000US-0226967P.
 PR 22-AUG-2000; 2000US-0226968P.
 PR 22-AUG-2000; 2000US-0226969P.
 PR 22-AUG-2000; 2000US-0226970P.
 PR 22-AUG-2000; 2000US-0226971P.
 PR 22-AUG-2000; 2000US-0226972P.
 PR 22-AUG-2000; 2000US-0226973P.
 PR 22-AUG-2000; 2000US-0226974P.
 PR 22-AUG-2000; 2000US-0226975P.
 PR 22-AUG-2000; 2000US-0226976P.
 PR 22-AUG-2000; 2000US-0226977P.
 PR 22-AUG-2000; 2000US-0226978P.
 PR 22-AUG-2000; 2000US-0226979P.
 PR 22-AUG-2000; 2000US-0226980P.
 PR 22-AUG-2000; 2000US-0226981P.
 PR 22-AUG-2000; 2000US-0226982P.
 PR 22-AUG-2000; 2000US-0226983P.
 PR 22-AUG-2000; 2000US-0226984P.
 PR 22-AUG-2000; 2000US-0226985P.
 PR 22-AUG-2000; 2000US-0226986P.
 PR 22-AUG-2000; 2000US-0226987P.
 PR 22-AUG-2000; 2000US-0226988P.
 PR 22-AUG-2000; 2000US-0226989P.
 PR 22-AUG-2000; 2000US-0226990P.
 PR 22-AUG-2000; 2000US-0226991P.
 PR 22-AUG-2000; 2000US-0226992P.
 PR 22-AUG-2000; 2000US-0226993P.
 PR 22-AUG-2000; 2000US-0226994P.
 PR 22-AUG-2000; 2000US-0226995P.
 PR 22-AUG-2000; 2000US-0226996P.
 PR 22-AUG-2000; 2000US-0226997P.
 PR 22-AUG-2000; 2000US-0226998P.
 PR 22-AUG-2000; 2000US-0226999P.
 PR 22-AUG-2000; 2000US

PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488783/53.
N-PSDB; AAS26175.

New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.

Claim 11; SEQ ID NO 1141; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
pathological condition. Antibodies to the proteins can also be used in
alleviating symptoms associated with the disorders and in diagnostic
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
(ELISA). Disorders which are diagnosed or treated include autoimmune
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac

CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 33.2%; Score 935; DB 4; Length 165;
Best Local Similarity 98.8%; Pred. No. 5.5e-57;
Matches 165; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 37 KRDDTKVSDTYKLPRLIEKKRRDRINECIAQKLDLLPEHLKLTTLGHLEKAVVLETLK 96
Db 1 KRDDTK--DTYKLPRLIEKKRRDRINECIAQKLDLLPEHLKLTTLGHLEKAVVLETLK 58

Qy 97 HLKALTALTEQQHQKIIALQNGERSLSKSPISQDLDAFHSGFQCAKEVLYLSRFESWTP 156
Db 59 HLKALTALTEQQHQKIIALQNGERSLSKSPISQDLDAFHSGFQCAKEVLYLSRFESWTP 118

Qy 157 REPRCVOLINHLHAVATQFLPTQLLTQQVPLSKGTGAPSAAGSAAA 203
Db 119 REPRCVOLINHLHAVATQFLPTQLLTQQVPLSKGTGAPSAAGSAAA 165

RESULT 14
ABUS5257
ID ABUS5257 standard; protein; 165 AA.
XX AC ABUS5257;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #344.
XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 28-JUN-2000; 2000US-0214856P.
XX PR 07-JUL-2000; 2000US-0216847P.
XX PR 07-JUL-2000; 2000US-0216890P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.

PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 03-SEP-2000; 2000US-0229509P.
PR 03-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26590.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1556; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
pathological condition. Antibodies to the proteins can also be used in
alleviating symptoms associated with the disorders and in diagnostic
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
(ELISA). Disorders which are diagnosed or treated include autoimmune
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodysplasia,
nervous system disorders e.g. Alzheimer's disease, infections caused by
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
and many other disorders listed in the specification. The polypeptides
can also be used to aid wound healing and epithelial cell proliferation,
to prevent skin aging due to sunburn, to maintain organs before
transplantation, for supporting cell culture of primary tissues, to
regenerate tissues and in chemotaxis. The polypeptides can also be used
as a food additive or preservative to increase or decrease storage
capabilities, fat content, lipid, protein, carbohydrate, vitamins,
minerals, cofactors and other nutritional components. The present
sequence represents a novel secreted protein of the invention. Note: The
sequence data for this patent did not form part of the printed

Query Match 29.8%; Score 750; DB 4; Length 150;
Best Local Similarity 97.4%; Pred. No. 2.le-50;
Matches 148; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY	37	KRDDTKVSDTYKLP	PHLEIEKKRRDRINECIAQIKOLLPEHLKLTTLTGLHLEKAVVLELT	96
Db				
QY	1	KRDDTK-	DTYKXHLRIEKKRRDRINECIAQIKOLLPEHLKLTTLTGLHLEKAVVLELT	58
Db				
QY	97	HLKALTALTEQQHQKII	ALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYLSRFESWTP	156
Db				
QY	59	HLKALTALTEQQHQKII	ALQNGERSLKSPIQSDLDFAHSGFQTCAKEVLQYLSRFESWTP	118
Db				
QY	157	REPRCVQLINHLH	AVATQFLPTPOLITQOVPL	188
Db				
QY	119	REPRCVQLINHLH	AVATQFLPTPOLITQOVPL	150
Db				

Search completed: December 15, 2004, 19:52:24
Job time : 163 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:13:29 ; Search time 698.631 Seconds
(without alignments)
11518.321 Million cell updates/sec

Title: US-10-078-650-11

Perfect score: 1511

Sequence: 1 catgagcaaggaattctctc.....gtgagaagttaataacct 1511

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1511	100.0	1511	4 AAF74776	AAF74776 Human DEC
2	1378	91.2	3521	10 ADC30675	ADC30675 Human nov
3	1378	91.2	3641	4 AAF74767	AAF74767 Human DEC
4	1378	91.2	3641	6 ABS76400	ABS76400 cDNA enco
5	1378	91.2	3641	9 ACC72416	ACC72416 Human ova
6	1378	91.2	3641	12 ADN05013	ADN05013 Antipsori
7	1339	88.6	1471	10 ACD19338	ACD19338 cDNA enco
8	1323	87.6	2304	6 ABT05457	ABT05457 DNA of NO
9	552	36.5	909	10 ABX74439	ABX74439 Human cDN
10	476	31.5	498	4 AAS26175	AAS26175 Human cDN
11	476	31.5	498	8 ABX73516	ABX73516 Human nov
12	366	24.2	454	4 AAS26590	AAS26590 Human cDN
13	366	24.2	454	8 ABX73931	ABX73931 Human nov
14	342	22.6	628	10 ACD19339	ACD19339 cDNA enco
15	274	18.1	285	4 AAS24787	AAS24787 Human ova
16	274	18.1	397	5 AAH83421	AAH83421 Human ova
17	128	8.5	502	3 AAC04136	AAC04136 Human sec
18	128	8.5	896	10 ADC32475	ADC32475 Human nov
19	54	3.6	1421	4 AAF74777	AAF74777 Mouse DEC
20	46	3.0	300	3 AAQ01182	AAQ01182 Human col
21	32	2.1	8147	6 ABS73313	ABS73313 DNA enco

C	22	32	2.1	8147	10	ADK11302	Adk11302 Human CRE
	23	31	2.1	1679	8	ABZ78139	Abz78139 Human can
	24	31	2.1	3060	2	AAX39667	Aax39667 Renal can
	25	31	2.1	4725	5	AAS76375	Aas76375 DNA enco
	26	31	2.1	4739	6	ABL63316	AbL63316 Breast ca
	27	31	2.1	4739	8	ACC00031	Acc00031 Human Bcr
	28	31	2.1	4739	10	ADG32836	Adg32836 Human DNA
	29	31	2.1	4756	5	AAS85030	Aas85030 DNA enco
	30	31	2.1	4775	5	AAS76377	Aas76377 DNA enco
	31	31	2.1	5795	5	AAS85031	Aas85031 DNA enco
	32	30	2.0	884	12	ADZ28379	Adz28379 Human KPP
	33	30	2.0	3544	5	AAS81565	Aas81565 DNA enco
	34	30	2.0	6335	6	ABK83722	Abk83722 Human cDN
	35	30	2.0	6335	11	ADN95256	Adn95256 Human BEC
	36	29	1.9	186	12	ADF77133	Adf77133 M08-TAG c
	37	29	1.9	201	12	ADF77129	Adf77129 E25-TAG c
	38	29	1.9	526	12	ACH78213	Ach78213 Human gen
	39	29	1.9	529	12	ACH77958	Ach77958 Human gen
	40	29	1.9	546	6	ABQ41782	Abq41782 Oligonuc1
	41	29	1.9	546	6	ABQ41783	Abq41783 Oligonuc1
	42	29	1.9	734	5	ADL46045	Adl46045 Human ova
	43	29	1.9	797	6	ABK34450	Abk34450 Human cDN
	44	29	1.9	1130	6	ABK92146	Abk92146 Prostate
	45	29	1.9	1130	12	ADP13302	Adp13302 Renal cel

ALIGNMENTS

RESULT 1
AAF74776
ID AAF74776 standard; cDNA; 1511 BP.
XX
AC AAF74776;
XX
DT 17-MAY-2001 (first entry)
XX
DE Human DEC2b encoding cDNA sequence SEQ ID NO:11.
XX
KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
KW basic helix loop helix protein; cell differentiation; proliferation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1456
FT /*tag= a
FT /product= "DEC2b"
FT /note= "bHLH type transcription factor"
XX
PN WO200114551-A1.
XX
PD 01-MAR-2001.
XX
PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-002333286.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Fujimoto K, Shin M, Kato Y;
DR WPI; 2001-202935/20.
DR P-PSDB; AAB70693.
XX
PT DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
PS Claim 1; Page 63-68; 83pp; Japanese.
XX
CC The present invention describes a basic helix loop helix (bHLH) type
transcription factor designated DEC2. DEC2 can be used as a tool in the

RESULT 3
 AAF74767
 ID AAF74767 standard; cDNA; 3641 BP.
 XX
 AC AAF74767;
 XX
 DT 17-MAY-2001 (first entry)
 XX
 DE Human DEC2a encoding cDNA sequence SEQ ID NO:1.
 XX
 KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
 KW basic helix loop helix protein; cell differentiation; proliferation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT Key 135..1583
 FT CDS /*tag= a
 FT /product= "DEC2a"
 FT /note= "bHLH type transcription factor"
 XX
 PN WO200114551-A1.
 XX
 XX 01-MAR-2001.
 XX
 XX 19-JUN-2000; 2000WO-JP003991.
 XX
 XX 19-AUG-1999; 99JP-00233286.
 PR (CHUS) CHUGAI SEIYAKU KK.
 XX
 XX Fujimoto K, Shin M, Kato Y;
 PI WPI; 2001-202935/20.
 DR P-PSDB; AAF70692.
 DR
 XX
 XX DEC2 is a basic helix loop helix protein of the DEC family for use in
 PT development of drugs for treatment of disorders of cell differentiation
 PT and proliferation.
 XX
 XX Claim 1; Page 48-55; 83pp; Japanese.
 PS
 CC The present invention describes a basic helix loop helix (bHLH) type
 CC transcription factor designated DEC2. DEC2 can be used as a tool in the
 CC development of drugs for the treatment and prevention of disorders
 CC involving cell differentiation and proliferation. The present sequence
 CC encodes the specifically claimed human DEC2a protein, as given in the
 CC present invention
 XX
 SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
 Query Match 91.2%; Score 1378; DB 4; Length 3641;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 134 GATACCTACAAATTACCGACAGATTATAGAAAAGAAAAGAGACCGAATTAATGAA 193
 DB 261 GATACCTACAAATTACCGACAGATTATAGAAAAGAAAAGAGACCGAATTAATGAA 320
 QY 194 TGCATTGCTCAGCTGAAAGATTACTGCTGAACATCTGAAATTGACAACTCTGGACAT 253
 DB 321 TGCATTGCTCAGCTGAAAGATTACTGCTGAACATCTGAAATTGACAACTCTGGACAT 380
 QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTGAAACACTTAAAGCTTTAACCGCCTTA 313
 DB 381 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTGAAACACTTAAAGCTTTAACCGCCTTA 440
 QY 314 ACCGACCAACAGCATCAGAAAGATTAATTGCTTTACAGAAATGGGAGCGATCTGAAATCG 373
 DB 441 ACCGACCAACAGCATCAGAAAGATTAATTGCTTTACAGAAATGGGAGCGATCTGAAATCG 500
 QY 374 CCCATTCACTCCGACTTGGATGGTTCCTCACTCGGGATTTCAAACATGCGCCAAAGAGTC 433

Db 501 CCCATTCACTCCGACTTGGATGGTTCCTCACTCGGGATTTCAAACATGCGCCAAAGAGTC 560
 QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTG 493
 Db 561 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTG 620
 QY 494 ATCAACCACTTGACCGCGTGGCCACCCAGTCTTCCGCCACCCCGCAGCTGTGTACTCAA 553
 Db 621 ATCAACCACTTGACCGCGTGGCCACCCAGTCTTCCGCCACCCCGCAGCTGTGTACTCAA 580
 QY 554 CAGGTCCCTCTGAGCAAAAGCACCAGCGCTCCCTCGGCCCGCGGGTCCCGGCGCGCCCC 613
 Db 681 CAGGTCCCTCTGAGCAAAAGCACCAGCGCTCCCTCGGCCCGCGGGTCCCGGCGCGCCCC 740
 QY 614 TGCCTGAGCGCGCGGGGCGAAGCTGGAGCCCTCGCTACTCGTGCCTCATCCAG 673
 Db 741 TGCCTGAGCGCGCGGGGCGAAGCTGGAGCCCTCGCTACTCGTGCCTCATCCAG 800
 QY 674 CGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACGACACACGAGCGGCTAC 733
 Db 801 CGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACGACACACGAGCGGCTAC 860
 QY 734 GCGCGGAAAGCGCGCGCGCGGACCGCGAGAAAGCAAGCGCGGGGCGGAGCGCGC 793
 Db 861 GCGCGGAAAGCGCGCGCGCGGACCGCGAGAAAGCAAGCGCGGGGCGGAGCGCGC 920
 QY 794 GTCACCATCAAGCAGGAGCTCCCGGGAGGACTCCCGCGCGCCCAAGAGGATGAAGCTG 853
 Db 921 GTCACCATCAAGCAGGAGCTCCCGGGAGGACTCCCGCGCGCCCAAGAGGATGAAGCTG 980
 QY 854 GATTCCCG 913
 Db 981 GATTCCCG 1040
 QY 914 GCGCTTCTGGGCG 973
 Db 1041 GCGCTTCTGGGCG 1100
 QY 974 CTCAGTCTCGTGTGGCGGTTTCGGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1033
 Db 1101 CTCAGTCTCGTGTGGCGGTTTCGGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1160
 QY 1034 GCG 1093
 Db 1161 GCG 1220
 QY 1094 GTGCAGCGCGCTTCTTGACAAAGAGCGCGCTTGAGAAAGTATCTGTACCCGCGCGCGGTGCC 1153
 Db 1221 GTGCAGCGCGCTTCTTGACAAAGAGCGCGCTTGAGAAAGTATCTGTACCCGCGCGCGGTGCC 1280
 QY 1154 GCGCGCGTTCGCTGTATACCCGCGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1213
 Db 1281 GCGCGCGTTCGCTGTATACCCGCGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1340
 QY 1214 GCGCGCGTTCG 1273
 Db 1341 GCGCGCGTTCG 1400
 QY 1274 CCGGAGAAAGCG 1333
 Db 1401 CCGGAGAAAGCG 1460
 QY 1334 GCGCGCGACCCCGACACCG 1393
 Db 1461 GCGCGCGACCCCGACACCG 1520
 QY 1394 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCGCAAGAAAGAGCTCCC 1453
 Db 1521 CCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCGCAAGAAAGAGCTCCC 1580
 QY 1454 TGAATCCTTGGTCCCGAAGGAGCGAGGTTCAAGCAGAGTGAGAGTTAAATACCTT 1511

Db	1221	GTGAGCCCTTCCTGGACAAGAGCGGCTCGAGAAGTATCTGTACCGCGGGCGCTGCC	1280
Qy	1154	GCCCGCTTCGCGTGTCTATACCCCGGCATCCCCCGCCCGCGGAGCGCGGGAGCGCGCC	1213
Db	1281	GCCCGCTTCGCGTGTCTATACCCCGGCATCCCCCGCCCGCGGAGCGCGGAGCGCGCC	1340
Qy	1214	GCGCGCGCTGCGCGCGCGCGCGCGGTCCTGCCCTGTCTCGGTGTGTGTGCGCCGCT	1273
Db	1341	GCGCGCGCTGCGCGCGCGCGCGCGGTCCTGCCCTGTCTCGGTGTGTGTGCGCCGCT	1400
Qy	1274	CCCGAGAAGCGGGCGCGCGCGCGACCTCTCGCGACGAGGTGGCGGCCCTTTGGG	1333
Db	1401	CCCGAGAAGCGGGCGCGCGCGCGACCTCTCGCGACGAGGTGGCGGCCCTTTGGG	1460
Qy	1334	GCGCGCACCCCHAGCACCGCGACGGCGGACCCACTGCGCTTCGCGCGGGCCCCGGCAG	1393
Db	1461	GCGCGCACCCCHAGCACCGCGACGGCGGACCCACTGCGCTTCGCGCGGGCCCCGGCAG	1520
Qy	1394	CCGGGGAAACCGGAGAGCTGTGCTCAGGAAGATCCCTCGCAGCCAGAAAGGAAGTCCC	1453
Db	1521	CCGGGGAAACCGGAGAGCTGTGCTCAGGAAGATCCCTCGCAGCCAGAAAGGAAGTCCC	1580
Qy	1454	TGAATCTCTGGTCCCGAAGACGGAGGTTCAAGCAGAGTGAGAAGTTAAAAACCTC	1511
Db	1581	TGAATCTCTGGTCCCGAAGACGGAGGTTCAAGCAGAGTGAGAAGTTAAAAACCTC	1638

RESIST. 6

RESULTS 6
ADN05013
ID ADN05013 standard; cDNA; 3641 BP.

AC ADN05013;

DT 01-JUL-2004 (first entry)

XX
DE Antipsoriatic cDNA sequence #721.

ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

OS Homo sapiens.

XX PN WO2004028479-A2.

XX PD 08-APR-2004.

XX
PE
25-SEP-2003. 2003WC-IIS030907XX
PB
25-SEP-2002. 2002US-0414006B

XX PA (GETH) GENTECH TNC

XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
PI

XX
DB
WPT: 2004-305105/28

DR WFL; 2004-305105/
DR P-PSDB: APN05014.

XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.

PS Claim 1: SEO ID NO 1407: 3069nn. English

xx The invention relates to novel polynucleotide and polypeptides for
 CC treating porrias or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating porrias in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.

Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;

Query Match 91.2%: score 1378: DB 12: Length 3641:

	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;	
	Matches	1378;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	134	GATACCTTACAAATTACCGCACAGATTAAATGAAAAAAGAAAAGAGACCAGAAATTAATGA	193							
DB	261	GATACCTTACAAATTACCGCACAGATTAAATGAAAAAAGAAAAGAGACCAGAAATTAATGA	320							
QY	194	TGCATTTCCTCAGCTGAAAGAATTACTGCCTGGAAACATCTGAAATTGACAACCTCTGGGACAT	253							
DB	321	TGCATTTCCTCAGCTGAAAGAATTACTGCCTGGAAACATCTGAAATTGACAACCTCTGGGACAT	380							
QY	254	CTGGAGAAAGCTGTAGTCTTTGGAAATAACTTTGAAACACATTTAAAGCTTTAAACCGCCTTA	313							
DB	381	CTGGAGAAAGCTGTAGTCTTTGGAAATAACTTTGAAACACATTTAAAGCTTTAAACCGCCTTA	440							
QY	314	ACGAGCAACAGCATCAGAAAGATAATTTGCTTTTACAGATGGGAGCGATCTCTGAAATCG	373							
DB	441	ACGAGCAACAGCATCAGAAAGATAATTTGCTTTTACAGATGGGAGCGATCTCTGAAATCG	500							
QY	374	CCATTTCAGTCGCAGCTTGGATGGTTCCTACTCGGATTTCAAACATGCGCCAAGAGTC	433							
DB	501	CCATTTCAGTCGCAGCTTGGATGGTTCCTACTCGGATTTCAAACATGCGCCAAGAGTC	560							
QY	434	TTGCAATACCTCTCCCGTTTTGAGAGCTGACACCCAGGGAGCGCGGTGTGTCCAGCTG	493							
DB	561	TTGCAATACCTCTCCCGTTTTGAGAGCTGACACCCAGGGAGCGCGGTGTGTCCAGCTG	620							
QY	494	ATCAACACATTTGACGCGCTGTGGCCACCCAGATTTGTGCCACCCCGCAGCTGTTGACTCAA	553							
DB	621	ATCAACACATTTGACGCGCTGTGGCCACCCAGATTTGTGCCACCCCGCAGCTGTTGACTCAA	680							
QY	554	CAGTCCCTCTGAGCAAAAGSCACCGCGCTCCCTCGCGCGCGCGGTTCGCGCGCGCCCC	613							
DB	681	CAGTCCCTCTGAGCAAAAGSCACCGCGCTCCCTCGCGCGCGCGGTTCGCGCGCGCCCC	740							
QY	614	TGCTTGAGCGCGCGGGCGAAGCTGGAGCCCTCGCTTACTGCTGTGCCGTATCCAG	673							
DB	741	TGCTTGAGCGCGCGGGCGAAGCTGGAGCCCTCGCTTACTGCTGTGCCGTATCCAG	800							
QY	674	CGGACTCAGCCCAGCGCGAGCTCGCGCGCGAAGCAGACGACACCGACAGGCTAC	733							
DB	801	CGGACTCAGCCCAGCGCGAGCTCGCGCGCGAAGCAGACGACACCGACAGGCTAC	860							
QY	734	GGCGGGGAAGCCGAGGCGCGCGCGAGCCGCGAGAAAGGCAAAAGCGCGGGCGAGCGCG	793							
DB	861	GGCGGGGAAGCCGAGGCGCGCGCGAGCCGCGAGAAAGGCAAAAGCGCGGGCGAGCGCG	920							
QY	794	GTCAACATCAAGCAGAGAGCTCCCGGGAGGACTCCGCGCGCCCAAGAGGATGAAGCTG	853							
DB	921	GTCAACATCAAGCAGAGAGCTCCCGGGAGGACTCCGCGCGCCCAAGAGGATGAAGCTG	980							
QY	854	GATTTCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	913							
DB	981	GATTTCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1040							
QY	914	GCGCCTTCTGGGCGCCGACCTGCGCGCGCGCGCGCGCTGCTGAGACCCGACGCGCGCTG	973							
DB	1041	GCGCCTTCTGGGCGCCGACCTGCGCGCGCGCGCGCGCTGCTGAGACCCGACGCGCGCTG	1100							
QY	974	CTCAGTCGCTGTGTGGCTTCTGCGTTCCTCTCGCTTCTCGAGTCGCGGCTAC	1033							
DB	1101	CTCAGTCGCTGTGTGGCTTCTGCGGAGCGGAGCGCGCTTCTCGAGTCGCGGCTAC	1160							
QY	1034	GC CGCGCGCCCTTCTGCGCTGCGCTTCTGCTTCTCTCGCTTCTCGAGTCGCGGCTAC	1093							
DB	1161	GC CGCGCGCCCTTCTGCGCTGCGCTTCTGCTTCTCTCGCTTCTCGAGTCGCGGCTAC	1220							
QY	1094	GTGAGGCCCTTCTTGACAAAGCGCGCTTGAGAAATATCTGTACCGCGCGCGGTGCC	1153							
DB	1221	GTGAGGCCCTTCTTGACAAAGCGCGCTTGAGAAATATCTGTGTACCGCGCGCGGTGCC	1280							
QY	1154	GC CC CGCTTCGCGCTGTATACCCCGCGCATCCCGCGCCCGCGCGCGCGCGCGCGCG	1213							

Db	1281	GC	CGCGTTC	CGCGTCT	ATAC	CCCGCGG	CATCC	CGCGCC	CGCGCGG	CAGCC	CGCGCAG	CGCGCC	1340			
Qy	1214	GC	CGCGCGT	CGCGCGCG	CGCGCGG	GTTC	CGCGT	CGCGT	CGCGT	CGCGT	CGCGT	CGCGCGT	1273			
Db	1341	GC	CGCGT	CGCGCG	CGCGCGG	GTTC	CGCGT	CGCGT	CGCGT	CGCGT	CGCGT	CGCGCGT	1400			
Qy	1274	CC	CGAGAA	CGCGG	CGCGCGG	CAC	CGTCT	CGCGC	CAC	CGAGT	CGCGCGC	CGTGG	1333			
Db	1401	CC	CGAAG	CGGCG	CGCGCG	CAC	CGTCT	CGCGC	CAC	CGAGT	CGCGCGC	CGTGG	1460			
Qy	1334	GC	CGCGCA	CCCC	CAGCA	CCCG	CAC	CGGCG	CAC	CCCA	CGTTC	CGCGCGC	CGCGAG	1393		
Db	1461	GC	CGCA	CCCC	CAGCA	CCCG	CAC	CGGCG	CAC	CCCA	CGTTC	CGCGCGC	CGCGAG	1520		
Qy	1394	CC	GGGNA	CCCGAG	CGTCT	CGT	CAG	AGAT	CCCT	CGCAG	CGCCAG	GAAGT	CGCC	1453		
Db	1521	CC	GGGNA	CCCGAG	CGTCT	CGT	CAG	AGAT	CCCT	CGCAG	CGCCAG	GAAGT	CGCC	1580		
Qy	1454	TG	AATC	CTTC	TCGTC	CCCGA	AGC	CGAGG	TTCA	AGCAG	AGT	GAGAA	GTAAAA	TAC	CGCT	1511
Db	1581	TG	AATC	CTTC	TCGTC	CCCGA	AGC	CGAGG	TTCA	AGCAG	AGT	GAGAA	GTAAAA	TAC	CGCT	1638

RESULT 7
ACD19338
ID ACD19338 standard: cDNA: 1471 BP.

DT 25-AUG-2003 (first entry)

XX
DE
cdNA encoding novel human protein #18.

Human; NOV; gene therapy; endocrine related disease; diabetes;
metabolism-related disease; obesity; central nervous system disorder;
Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
schizophrenia; depression; autoimmune disorder; inflammatory disorder;
psoriasis; allergy; lupus erythematosus; asthma; cancer;
inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
lung disease; emphysema; obstructive pulmonary disease; haemophilia;
stroke; infection; gene; ss.

OS Homo sapiens.

XX
PN
WC2003023002-A2.XX
PD
30-MAR-2003

XX

X

PR 07-SEP-2001;
PR 07-SEP-2001;

PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.

PR 17-SEP-2001; 2001US-0322781P.

PR 17-SEP-2001; 2001US-0322817P.

PR 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001: 2001US-0323636P.
PR 25-SEP-2001: 2001US-0324969P.

PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.

PR 17-APR-2002; 2002US-0373212P.

[illegible]XX
FA (CURA[®]) CURAGEN CORP.

PI : Spytek KA, Patturajan M, Go
PI Gerlach VL: Vernet CAM, Ell

PI Shimkets RA, Leach MD, Catt

PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX WPI; 2003-313242/30.
DR P-PSDB; ABO14645.
DR

xx New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
 PT and polynucleotides, useful in gene therapy, e.g. for treating or
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
 PT stroke or infections.

PS Claim 20: Page 125; 586pp; English.

The invention describes a new isolated polypeptide (NOVX). The NOVX polypeptide, nucleic acid and antibody are useful as therapeutics, particularly in the manufacture of a medicament for treating a syndrome associated with a human disease, which includes a pathology associated with NOVX polypeptide. The DNA encoding the protein is useful in gene therapy for treating the disease or condition. In particular, the NOVX polypeptide or polynucleotide is useful for treating endocrine/metabolism-related diseases (e.g. obesity or diabetes), central nervous system disorders (e.g. Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune and inflammatory disorders (e.g. psoriasis, allergy, lupus, erythematosis, asthma, inflammatory bowel disease, rheumatoid arthritis or osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian, prostate or brain cancers, or melanoma), liver diseases (e.g. liver cirrhosis), lung diseases (emphysema or obstructive pulmonary disease), haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic). These are also useful in developing powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications, and for monitoring the effects of drugs during clinical trials. This sequence encodes a novel human NOVX protein

XX SO Sequence 1471 BP; 288 A; 515 C; 436 G; 232 T; 0 U; 0 Other;

Query Match 88.6%; Score 1339; DB 10; Length 1471;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

134	QY	GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACACCGAATTAATGAA	133
133	DB	GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACACCGAATTAATGAA	192
194	QY	TGCATTGCTCAGCTGAAAGATTACTGCCTGAACTCTGAAATTTGACACTCTGGACAT	253
193	DB	TGCATTGCTCAGCTGAAAGATTACTGCCTGAACTCTGAAATTTGACACTCTGGACAT	252
254	QY	CTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTTGAAAACATTAAAGCTTTAACCGCCTTA	313
253	DB	CTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTTGAAAACATTAAAGCTTTAACCGCCTTA	312
314	QY	ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGGAGCGACTCTCTGAAATCG	373
313	DB	ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATGGGGAGCGACTCTCTGAAATCG	372
374	QY	CCCAATTCAGTCCGACTTGGATGCGTTCCTACTCGGGAATTTCAAACATGCGGCCAAAGAATC	433
373	DB	CCCAATTCAGTCCGACTTGGATGCGTTCCTACTCGGGAATTTCAAACATGCGGCCAAAGAATC	432
434	QY	TTGCAATACCTCTCCGGTTTGGAGACTGAGACACCCAGAGGACGCGGTGTGTCAGCTG	493
433	DB	TTGCAATACCTCTCCGGTTTGGAGACTGAGACACCCAGAGGACGCGGTGTGTCAGCTG	492
494	QY	ATCAACCACTTGCACGCGGTGGCCACCCAGTTCTTCCCAACCCCGAGCTGTTGACTCAA	553
493	DB	ATCAACCACTTGCACGCGGTGGCCACCCAGTTCTTCCCAACCCCGAGCTGTTGACTCAA	552
554	QY	CAGGTCCCTCTGAGCAAGGACCGGAGTCCCTCGGCGCGCGGGTCCGCGGCCGCCCCC	613
553	DB	CAGGTCCCTCTGAGCAAGGACCGGAGTCCCTCGGCGCGCGGGTCCGCGGCCGCCCCC	612
614	QY	TGCCTGGAGCGCGCGGGGAGAGCTGGAGCCCTCGCCTACTTCGCTGCCGTCATCCAG	673

Db 613 TGCTGAGCGCGGGGCGAAGCTGGAGCCCTCGCTACTGCTGCTGCTATCCAG 672
 Qy 674 CGGACTCAGCCGAGCGCGAGCTGCGCGCGGAGAAACGACGACACCGACAGCGCTAC 733
 Db 673 CGGACTCAGCCGAGCGCGAGCTGCGCGCGGAGAAACGACGACACCGACAGCGCTAC 732
 Qy 734 GGGCGGAAACCGAGCGCGCGGAGCGCGGAGAAAGGAAAGCGCGGGGGGAGCGCC 793
 Db 733 GGGCGGAAACCGAGCGCGCGGAGCGCGGAGAAAGGAAAGCGCGGGGGGAGCGCC 792
 Qy 794 GTCACCATCAGCAGGAGCTCCCGGGAGGAACTCCCGCGCGCCAAAGAGGATGAAGCTG 853
 Db 793 GTCACCATCAGCAGGAGCTCCCGGGAGGAACTCCCGCGCGCCAAAGAGGATGAAGCTG 852
 Qy 854 GATTCCCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 913
 Db 853 GATTCCCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912
 Qy 914 GCGCTTCTGGGGCCCGACCTCGCGCGCGCGCGCGCGCTGTGAGACCCGACGCGCGCTG 973
 Db 913 GCGCTTCTGGGGCCCGACCTCGCGCGCGCGCGCGCGCTGTGAGACCCGACGCGCGCTG 972
 Qy 974 CTGAGCTCGCTGTGGGCTTGGCGGAGGCGGAGGCGCGCGCTTCCCGAGCGCGCGCG 1033
 Db 973 CTGAGCTCGCTGTGGGCTTGGCGGAGGCGGAGGCGCGCGCTTCCCGAGCGCGCGCG 1032
 Qy 1034 GCGCGCGCGCGCTTGCCTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
 Db 1033 GCGCGCGCGCGCTTGCCTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTG 1092
 Qy 1094 GTGAGCGCTTCTGAGCAAGAGCGCGCTGGAGAGTATCTGTACCCGCGCGCGGTGCC 1153
 Db 1093 GTGAGCGCTTCTGAGCAAGAGCGCGCTGGAGAGTATCTGTACCCGCGCGCGGTGCC 1152
 Qy 1154 GCGCGCTTCCGCTGTATACCCCGCGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1213
 Db 1153 GCGCGCTTCCGCTGTATACCCCGCGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1212
 Qy 1214 GCGCGCGCTGCT 1273
 Db 1213 GCGCGCGCTGCT 1272
 Qy 1274 CCGGAGAGCG 1333
 Db 1273 CCGGAGAGCG 1332
 Qy 1334 GCGCGCGCACCCCGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1393
 Db 1333 GCGCGCGCACCCCGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1392
 Qy 1394 CCGGGGAAACCGGAGAGCTTGTCTAGGAGATCCCTCGAGCGAGAAAGAAAGTCCG 1453
 Db 1393 CCGGGGAAACCGGAGAGCTTGTCTAGGAGATCCCTCGAGCGAGAAAGAAAGTCCG 1452
 Qy 1454 TGAATCTTGGTCCCGAA 1472
 Db 1453 TGAATCTTGGTCCCGAA 1471

RESULT 8
 ID ABT05457
 XX ABT05457 standard; DNA; 2304 BP.
 XX
 AC ABT05457;
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE DNA of NOVX 5 SEQ ID No 9.
 XX
 KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipemic;
 KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
 KW tranquilliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;

anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
 metabolic disorder; obesity; infectious disease; Alzheimer's disease;
 anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
 immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
 metabolic syndrome X; wasting disorder; cancer; neurological disorder;
 epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
 vesicular transport; cystic fibrosis; gastrointestinal disorder;
 diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
 multiple sclerosis; rheumatoid arthritis; transgenic animal;
 gene therapy; gene; ds.

Unidentified.

WO200246409-A2.

13-JUN-2002.

06-DEC-2001; 2001WO-US046586.

06-DEC-2000; 2000US-0251660P.

12-DEC-2000; 2000US-0255029P.

08-JAN-2001; 2001US-0280386P.

24-JAN-2001; 2001US-0258800P.

20-FEB-2001; 2001US-0269942P.

24-APR-2001; 2001US-0286183P.

20-AUG-2001; 2001US-0313627P.

12-SEP-2001; 2001US-0318712P.

(CURA-) CURAGEN CORP.

Guo X, Li L, Patturajan M, Shinkets RA, Casman SU, Malyankar UM;

Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;

Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;

Colman SD, Bisen AJ, Liu X, Padigaru M, Spaderna SK, Zerhusen BD;

WPI; 2002-547774/58.

P-PsDB; ABJ04644.

Novel isolated polypeptide, designated NOVX, useful for treating or preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and metabolic, neurodegenerative, immune and hematopoietic disorders.

Claim 9; Page 43-44; 421pp; English.

The invention relates to an isolated polypeptide, designated NOVX, comprising a sequence fully defined in the specification. The isolated protein, its encoding polynucleotide or an antibody created from the protein is useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder, or for treating or preventing a NOVX-associated disorder in a subject, preferably human. The isolated protein, its encoding polynucleotide or an antibody created from the protein are also useful for treating or preventing metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder, immune disorders, haematopoietic disorders, and various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X, wasting disorders associated with chronic diseases, and cancer. The isolated protein, its encoding polynucleotide or an antibody created from the protein are useful for treating or preventing neurological disorders such as epilepsy, stroke, mental disorders including mood, anxiety, schizophrenic disorders, disorders of vesicular transport such as cystic fibrosis, diabetes mellitus, goiter, gastrointestinal disorders including ulcerative colitis, other conditions associated with abnormal vesicle trafficking including AIDS, allergic reactions, multiple sclerosis and rheumatoid arthritis. A cell comprising the vector of the invention is useful for producing non-human transgenic animals. The polynucleotide of the invention can be used to treat disorders by gene therapy. This polynucleotide sequence represents the DNA encoding one of the isolated NOVX proteins of the invention

Sequence 2304 BP; 468 A; 784 C; 689 G; 363 T; 0 U; 0 Other;

Query Match		87.6%;	Score 1323;	DB 6;	Length 2304;	
Best Local Similarity		100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 1323;		Conservative	0;	Mismatches	0;	Indels
Qy	134	GATACCTCAAAATTTACCGCACGATTAATAGAAAAGAAAGAACGACCGAATTAATGAA	193			
Db	982	GATACCTCAAAATTTACCGCACGATTAATAGAAAAGAAAGAACGACCGAATTAATGAA	1041			
Qy	194	TGATTCCTCAGTGAAGATTTACTGCTGAACATCTGAATTCGAATTCGAGCAT	253			
Db	1042	TGATTCCTCAGTGAAGATTTACTGCTGAACATCTGAATTCGAATTCGAGCAT	1101			
Qy	254	CTGGAGAAAGCTGTAGTCTTTGGAATTAATCTTTGAACACCTTAAAGCTTTAAACCCCTTA	313			
Db	1102	CTGGAGAAAGCTGTAGTCTTTGGAATTAATCTTTGAACACCTTAAAGCTTTAAACCCCTTA	1161			
Qy	314	ACCGAGCAACAGCATCAGAAAGATTAATGCTTTACAGAAATGGGAGCGATCTCTGAATCG	373			
Db	1162	ACCGAGCAACAGCATCAGAAAGATTAATGCTTTACAGAAATGGGAGCGATCTCTGAATCG	1221			
Qy	374	CCATTACGTCGACCTTGGATGGTTCCACTCGGATTTCAACATGCGCCAAAGAGTC	433			
Db	1222	CCATTACGTCGACCTTGGATGGTTCCACTCGGATTTCAACATGCGCCAAAGAGTC	1281			
Qy	434	TTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGTGTGTCCAGCTG	493			
Db	1282	TTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGTGTGTCCAGCTG	1341			
Qy	494	ATCAACCACTTGCACCGCTGGCCACCCAGTCTTTCGCCACCCCGCAGCTGTGACTCA	553			
Db	1342	ATCAACCACTTGCACCGCTGGCCACCCAGTCTTTCGCCACCCCGCAGCTGTGACTCA	1401			
Qy	554	CAGTCTCCTCTGAGCAAAAGCAGCGGCTCCTCGCGCGCGCGGFTCCCGCGCCGCC	613			
Db	1402	CAGTCTCCTCTGAGCAAAAGCAGCGGCTCCTCGCGCGCGCGGFTCCCGCGCCGCC	1461			
Qy	614	TGCTTGAGCGCGCGGCGGAGAGCTGGAGCCCTCTGCTACTGCTGCTCCGCTCATCAG	673			
Db	1462	TGCTTGAGCGCGCGGCGGAGAGCTGGAGCCCTCTGCTACTGCTGCTCCGCTCATCAG	1521			
Qy	674	CGACTCAGCCAGCGCGGAGCTCGCGCGGAGAGCAGACACCGACACCGCGCTAC	733			
Db	1522	CGACTCAGCCAGCGCGGAGCTCGCGCGGAGAGCAGACACCGACACCGCGCTAC	1581			
Qy	734	GGCGGGAAGCGAGGCGCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	793			
Db	1582	GGCGGGAAGCGAGGCGCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1641			
Qy	794	GTACCATCAAGCAGGAGCTCCGCGGAGAGCTCGCGCGGCGGCGGCGGCGGCGGCG	853			
Db	1642	GTACCATCAAGCAGGAGCTCCGCGGAGAGCTCGCGCGGCGGCGGCGGCGGCGGCG	1701			
Qy	854	GATTCG	913			
Db	1702	GATTCG	1761			
Qy	914	CGCTTCTGGGCG	973			
Db	1762	CGCTTCTGGGCG	1821			
Qy	974	CTCAGCTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1033			
Db	1822	CTCAGCTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1881			
Qy	1034	GCCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1093			
Db	1882	GCCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1941			
Qy	1094	GTGAGCGCTTCTGGAAGAGCGGCTGAGAGATATCTGATCCGCGCGCGCGCTGCG	1153			
Db	1942	GTGAGCGCTTCTGGAAGAGCGGCTGAGAGATATCTGATCCGCGCGCGCGCTGCG	2001			
Qy	1154	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1213			

Db	2002	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2061			
Qy	1214	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1273			
Db	2062	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2121			
Qy	1274	CCCGAGAGCG	1333			
Db	2122	CCCGAGAGCG	2181			
Qy	1334	GGCGCGCACCCCGAGACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1393			
Db	2182	GGCGCGCACCCCGAGACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2241			
Qy	1394	CCCGGGAACCCGAGAGCTGCTGCTAGGAAGATCCCTCGCGCGCGCGCGCGCGCGCG	1453			
Db	2242	CCCGGGAACCCGAGAGCTGCTGCTAGGAAGATCCCTCGCGCGCGCGCGCGCGCGCG	2301			
Qy	1454	TGA 1456				
Db	2302	TGA 2304				
RESULT 9						
ABX74439						
ID	ABX74439	standard;	cdna;	909	BP.	
XX	AC	ABX74439;				
XX	DT	21-MAR-2003	(first entry)			
XX	DE	Human	cdna	sequence #11	up-regulated in CC-RCC patients.	
XX	KW	Human;	microarray;	solid surface;	immobilised probe; CC-RCC;	
XX	KW	differential	expression profile;	aggressive	CC-RCC tumour type;	
XX	KW	non-aggressive	CC-RCC tumour type;	clear cell	renal carcinoma;	
XX	OS	gene	expression profiling;	tumour	tissue; gene; ss.	
XX	PN	Homo	sapiens.			
XX	PD	WO200279411-A2.				
XX	PF	10-OCT-2002.				
XX	PR	29-MAR-2002;	2002WO-US009576.			
XX	PR	29-MAR-2001;	2001US-0279411P.			
XX	PA	(VAND-)	VAN ANDEL INST.			
XX	PI	Haab B,	Rhodes D,	Teh BT,	Takashi M;	
XX	DR	WPI;	2003-040679/03.			
XX	PT	New	microarray,	comprising	a matrix of cdna	probe from a set of probes
XX	PT	immobilized	to a solid surface	in predetermined	order, useful in the	prognosis of patients with clear cell renal carcinoma.
XX	PS	Claim	35;	Page 119-120;	179pp;	English.
XX	CC	The present invention relates to a microarray comprising a matrix of at least one cdna probe from a set of probes immobilised to a solid surface in a predetermined order, where a row of pixels corresponds to replicates of one distinct probe from the set. The probes are complementary to nucleic acid sequences that are expressed differentially in aggressive as compared to non-aggressive types of clear cell renal carcinoma (CC-RCC) and that hybridise to the probes under high stringency conditions. The microarray is useful for the prognosis of patients with CC-RCC, wherein aggressive and non-aggressive CC-RCC tumour types are characterised by differential expression profiles of genes that hybridise with one or more probes immobilised on the microarray. The arrays are useful for gene expression profiling of tumour and normal tissues. The present sequence				

CC represents a human cDNA sequence up-regulated in CC-RCC patients

```
XX XX
SQ Sequence 909 BP; 218 A; 292 C; 256 G; 143 T; 0 U; 0 Other;

Query Match          36.5%; Score 552; DB 10; Length 909;
Best Local Similarity 100.0%; Pred. No. 8.5e-243;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTCGAATTAACCTTTGAAACACTTAAAGCT 301
Db 181 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTCGAATTAACCTTTGAAACACTTAAAGCT 240

QY 302 TTAACCGCTTAAACCGAGCAACAGCATCAGAACATTAATTCCTTACAGATGGGAGCA 361
Db 241 TTAACCGCTTAAACCGAGCAACAGCATCAGAACATTAATTCCTTACAGATGGGAGCA 300

QY 362 TCTCTGAATTCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAAAATGC 421
Db 301 TCTCTGAATTCGCCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAAAATGC 360

QY 422 GCCAAGAGTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCGGG 481
Db 361 GCCAAGAGTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCGGG 420

QY 482 TGTGTCCAGTGTATCAACACATTCGACGCGCTGGCCACCCAGTTCCTTGCCACCCCGCAG 541
Db 421 TGTGTCCAGTGTATCAACACATTCGACGCGCTGGCCACCCAGTTCCTTGCCACCCCGCAG 480

QY 542 CTGTTGACTCAACAGTCCCTCTGAGAAAGGACCGCGCTCCCTCGCGCCCGGGTCC 601
Db 481 CTGTTGACTCAACAGTCCCTCTGAGAAAGGACCGCGCTCCCTCGCGCCCGGGTCC 540

QY 602 GCGGCGCGCCCTCGCTGAGCGCGCGGAGAGCTGGAGCCCTCGCTACTGCGTG 661
Db 541 GCGGCGCGCCCTCGCTGAGCGCGCGGAGAGCTGGAGCCCTCGCTACTGCGTG 600

QY 662 CCGCTCATCAGCGGACTCAGCCACGCGGAGCTCGCCCGGAGAACGACGACAC 721
Db 601 CCGCTCATCAGCGGACTCAGCCACGCGGAGCTCGCCCGGAGAACGACGACAC 660

QY 722 GACAGCGGTACGCGCGGAGCGCGGCGCGGAGCGGAGCGGAGCGGAGCGG 781
Db 661 GACAGCGGTACGCGCGGAGCGCGGCGCGGAGCGGAGCGGAGCGGAGCGG 720

QY 782 GGGCGGAGCGCG 793
Db 721 GGGCGGAGCGCG 732

RESULT 10
AAS26175
ID AAS26175 standard; cDNA; 498 BP.
XX AC
AC AAS26175;
XX AC
XX DT 07-NOV-2001 (first entry)
XX DE
DE Human cDNA encoding a novel secreted protein, Seq ID 354.
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS
OS Homo sapiens.
XX PN
PN WO200155322-A2.
XX XX
```

PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239353P.
 PR 13-OCT-2000; 2000US-0239357P.
 PR 20-OCT-2000; 2000US-0240360P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244517P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0251799P.
 PR 06-DEC-2000; 2000US-0251799P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 03-JAN-2001; 2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX ABX73516
 XX WPI; 2001-488783/53.
 DR

DR P-PSDE; AAUI6183.
 XX New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX Claim 1; SEQ ID NO 354; 980pp; English.
 XX The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Antibodies to the proteins can also be used in
 CC alleviating symptoms associated with the disorders and in diagnostic
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 CC and many other disorders listed in the specification. The polypeptides
 CC can also be used to aid wound healing and epithelial cell proliferation,
 CC to prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence encodes a novel secreted protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed

Query Match 31.5%; Score 476; DB 4; Length 498;
 Best Local Similarity 100.0%; Pred. No. 6.4e-208;
 Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 193
 Db 21 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 80
 QY 194 TGCATTGCTCAGCTGAAAGATTTACTGCTGAACTCTGAAATGACACTCTGGACAT 253
 Db 81 TGCATTGCTCAGCTGAAAGATTTACTGCTGAACTCTGAAATGACACTCTGGACAT 140
 QY 254 CTGGAGAAAGCTGTAGTCTTGGATTAACTTTGAAACACTTAAAGCTTTAACCGCTTA 313
 Db 141 CTGGAGAAAGCTGTAGTCTTGGATTAACTTTGAAACACTTAAAGCTTTAACCGCTTA 200
 QY 314 ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 373
 Db 201 ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 260
 QY 374 CCNATTGAGTCCGACTTGGATGGTTCCTCCACTCGGATTTCAAACATGCGCAAGAGTC 433
 Db 261 CCCATTGAGTCCGACTTGGATGGTTCCTCCACTCGGATTTCAAACATGCGCAAGAGTC 320
 QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTCTCCAGCTG 493
 Db 321 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTCTCCAGCTG 380
 QY 494 ATCAACCACTTGACACGCGGTGGCCACCCAGTCTTGGCCACCCCGGAGCTGTGATCAA 553
 Db 381 ATCAACCACTTGACACGCGGTGGCCACCCAGTCTTGGCCACCCCGGAGCTGTGATCAA 440
 QY 554 CAGTCCCTCTGAGCAAAAGGACCGCGCTCCCTCGGCGCGCGGTCCCGGCGCGC 609
 Db 441 CAGTCCCTCTGAGCAAAAGGACCGCGCTCCCTCGGCGCGCGGTCCCGGCGCGC 496

RESULT 11
 ABX73516
 ID ABX73516 standard; DNA; 498 Bp.

XX AC ABX73516;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polynucleotide #344.
XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; aniallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216447P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR P-PSDB; ABU55256.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular, or
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 354; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention.
XX
SQ Sequence 498 BP; 142 A; 137 C; 115 G; 104 T; 0 U; 0 Other;

Query Match 31.5%; Score 476; DB 8; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.4e-208;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGCAATTAATGAA 193
DB 21 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGCAATTAATGAA 80
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCTGAAACATCTGAAATTCACAACCTCTCGGACAT 253
DB 81 TGCATTGCTCAGCTGAAAGATTACTGCTGAAACATCTGAAATTCACAACCTCTCGGACAT 140
QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACTTAAAGCTTTAACCGCCTTA 313
DB 141 CTGGAGAAAGCTGTAGTCTTGGAAATTAACTTTGAAACACTTAAAGCTTTAACCGCCTTA 200
QY 314 ACCGAGCAACAGCATCAGAGATTAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 373
DB 201 ACCGAGCAACAGCATCAGAGATTAATTGCTTTACAGAAATGGGAGCGATCTCTGAAATCG 260
QY 374 CCCATTGAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAACATCGCCCAAGAGTC 433
DB 261 CCCATTGAGTCCGACTTGGATGCGTTCCACTCGGGATTTCAACATCGCCCAAGAGTC 320
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCAGCTG 493
DB 321 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCAGCTG 380
QY 494 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTTGACTCAA 553
DB 381 ATCAACCACTTGACGCGCGTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTTGACTCAA 440
QY 554 CAGTCCCTCTGAGCAAAAGGACCGCGGCTCCCTCGGCGCGCGGTCCGCGCGCGC 609
DB 441 CAGTCCCTCTGAGCAAAAGGACCGCGGCTCCCTCGGCGCGCGGTCCGCGCGCGC 496

RESULT 12

AAAS26590	08-SEP-2000;	2000US-02312443P
ID	AAAS26590 standard; cDNA; 454 BP.	PR 08-SEP-2000;
XX	XX	PR 08-SEP-2000;
XX	AAAS26590;	PR 08-SEP-2000;
XX	XX	PR 08-SEP-2000;
XX	07-NOV-2001 (first entry)	PR 08-SEP-2000;
DT	Human cDNA encoding a novel secreted protein, Seq ID 769.	PR 14-SEP-2000;
XX	XX	PR 14-SEP-2000;
XX	Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;	PR 14-SEP-2000;
XX	Human; vasotropic; cerebroprotective; nootropic; neuroprotective;	PR 14-SEP-2000;
KW	cardiant; antibacterial; virucide; fungicide; ophthalmological; vulnerary;	PR 14-SEP-2000;
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;	PR 14-SEP-2000;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;	PR 14-SEP-2000;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;	PR 14-SEP-2000;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;	PR 21-SEP-2000;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;	PR 21-SEP-2000;
KW	cerebral ischaemia; angiogenesis; nervous system disorder;	PR 21-SEP-2000;
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;	PR 25-SEP-2000;
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;	PR 25-SEP-2000;
KW	preservative; antiproliferative.	PR 26-SEP-2000;
XX	XX	PR 26-SEP-2000;
XX	Homo sapiens.	PR 27-SEP-2000;
OS	XX	PR 27-SEP-2000;
XX	XX	PR 27-SEP-2000;
XX	WO200155322-A2.	PR 29-SEP-2000;
PN	XX	PR 29-SEP-2000;
XX	XX	PR 29-SEP-2000;
PD	02-AUG-2001.	PR 29-SEP-2000;
XX	XX	PR 29-SEP-2000;
XX	17-JAN-2001; 2001WO-US0001341.	PR 02-OCT-2000;
PF	XX	PR 02-OCT-2000;
XX	31-JAN-2000; 2000US-0179065P.	PR 02-OCT-2000;
PR	04-FEB-2000; 2000US-0180628P.	PR 02-OCT-2000;
PR	24-FEB-2000; 2000US-0184664P.	PR 02-OCT-2000;
PR	02-MAR-2000; 2000US-0186350P.	PR 13-OCT-2000;
PR	16-MAR-2000; 2000US-0189874P.	PR 13-OCT-2000;
PR	17-MAR-2000; 2000US-0190076P.	PR 13-OCT-2000;
PR	18-APR-2000; 2000US-0198123P.	PR 20-OCT-2000;
PR	19-MAY-2000; 2000US-0205515P.	PR 20-OCT-2000;
PR	07-JUN-2000; 2000US-0209467P.	PR 20-OCT-2000;
PR	28-JUN-2000; 2000US-0214886P.	PR 20-OCT-2000;
PR	30-JUN-2000; 2000US-0215135P.	PR 20-OCT-2000;
PR	07-JUL-2000; 2000US-0216647P.	PR 20-OCT-2000;
PR	07-JUL-2000; 2000US-0216800P.	PR 20-OCT-2000;
PR	11-JUL-2000; 2000US-0217487P.	PR 20-OCT-2000;
PR	11-JUL-2000; 2000US-0217496P.	PR 20-OCT-2000;
PR	14-JUL-2000; 2000US-0218290P.	PR 20-OCT-2000;
PR	26-JUL-2000; 2000US-0220963P.	PR 20-OCT-2000;
PR	26-JUL-2000; 2000US-0220964P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0224518P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0224519P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225213P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225214P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225266P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225267P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225268P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225270P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225447P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225472P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225758P.	PR 20-OCT-2000;
PR	14-AUG-2000; 2000US-0225759P.	PR 20-OCT-2000;
PR	18-AUG-2000; 2000US-0226279P.	PR 20-OCT-2000;
PR	22-AUG-2000; 2000US-0226661P.	PR 20-OCT-2000;
PR	22-AUG-2000; 2000US-0226688P.	PR 20-OCT-2000;
PR	22-AUG-2000; 2000US-0227182P.	PR 20-OCT-2000;
PR	23-AUG-2000; 2000US-0227009P.	PR 20-OCT-2000;
PR	30-AUG-2000; 2000US-0228924P.	PR 20-OCT-2000;
PR	01-SEP-2000; 2000US-0229287P.	PR 20-OCT-2000;
PR	01-SEP-2000; 2000US-0229343P.	PR 20-OCT-2000;
PR	01-SEP-2000; 2000US-0229344P.	PR 20-OCT-2000;
PR	01-SEP-2000; 2000US-0229345P.	PR 20-OCT-2000;
PR	03-SEP-2000; 2000US-0229509P.	PR 20-OCT-2000;
PR	05-SEP-2000; 2000US-0229513P.	PR 20-OCT-2000;
PR	06-SEP-2000; 2000US-0230437P.	PR 20-OCT-2000;
PR	06-SEP-2000; 2000US-0230438P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0231242P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0231243P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0249264P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0249265P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0249266P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-0249267P.	PR 20-OCT-2000;
PR	08-SEP-2000; 2000US-02	

29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 29-SEP-2000; 2000US-0236802P.
 29-SEP-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.
 13-OCT-2000; 2000US-0239935P.
 20-OCT-2000; 2000US-0240960P.
 20-OCT-2000; 2000US-0241785P.
 20-OCT-2000; 2000US-0241809P.
 01-NOV-2000; 2000US-0244617P.
 17-NOV-2000; 2000US-0249299P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 (ROSE/) ROSEN C A.
 (RUBE/) RUBEN S M.
 (BARA/) BARASH S C.
 Rosen CA, Ruben SM, Barash SC;
 WPI; 2003-147444/14.
 P-PSDB; ABUS5671.
 New polypeptides and nucleic acids, useful in gene therapy for treating,
 inhibiting or preventing e.g. neural, immune system, muscular,
 respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 renal disorders.
 Claim 1; SEQ ID NO 769; 402bp; English.
 The invention relates to human novel polypeptides and their associated
 polynucleotides. The polypeptides and polynucleotides are useful in gene
 therapy for treating, inhibiting or preventing neural disorders, immune
 system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left
 heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 leukemia), inflammatory diseases (e.g. septic shock, bursitis and
 appendicitis), allergic reactions and conditions (e.g. asthma), blood
 related disorders (e.g. thrombosis, atherosclerosis and myocardial
 infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
 human novel polynucleotides of the invention
 Query Match 24.2%; Score 366; DB 8; Length 454;
 Best Local Similarity 99.8%; Pred. No. 1.9e-157;
 Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 151 GCACAGATTATGAAAGAAAGACAGACCGAATTAATGATGCTTCAGCTGAA 210
 38 GCACAGATTATGAAAGAAAGACAGACCGAATTAATGATGCTTCAGCTGAA 97
 211 AGATTACTGCTGAAATCTGAAATGACAACTCTGGGACATCTGGGAGAAAGCTGTAGT 270
 98 AGATTACTGCTGAAATCTGAAATGACAWCTCTGGGACATCTGGGAGAAAGCTGTAGT 157
 271 CTGGGAATTAACCTTGAACACATTAAGCTTTAACCGCTTAACCGACACACATCA 330
 158 CTGGGAATTAACCTTGAACACATTAAGCTTTAACCGCTTAACCGACACACATCA 217
 331 GAAGATAATGCTTTACAGAAATGGGAGCGATCTCTGAATGCCCATTCAGTCCGACTT 390
 218 GAAGATAATGCTTTACAGAAATGGGAGCGATCTCTGAATGCCCATTCAGTCCGACTT 277

QY 391 GGATGGTTCCTCACTCGGGATTTTCAAAATGCGCCAAAGAGTCTTGCATATCTCTCCG 450
 DB 278 GGATGGTTCCTCACTCGGGATTTTCAAAATGCGCCAAAGAGTCTTGCATATCTCTCCG 337
 QY 451 GTTGTGAGCTGACACACCGAGGAGCGGGTGTCTCCAGCTGATCAACCACTTGCACGC 510
 DB 338 GTTGTGAGCTGACACACCGAGGAGCGGGTGTCTCCAGCTGATCAACCACTTGCACGC 397
 QY 511 CGTGGCCACCCAGTCTTGTCCACCCCGCAGCTGTGTGACTCAACAGTCCCTCTGAG 567
 DB 398 CGTGGCCACCCAGTCTTGTCCACCCCGCAGCTGTGTGACTCAACAGTCCCTCTGAG 454
 RESULT 14
 ACD19339
 ID ACD19339 standard; cDNA; 628 BP.
 AC ACD19339;
 XX
 DT 25-AUG-2003 (first entry)
 XX
 DB cDNA encoding novel human protein #19.
 XX
 KW Human; NOV; gene therapy; endocrine related disease; diabetes;
 KW metabolism-related disease; obesity; central nervous system disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003023002-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 09-SEP-2002; 2002WO-US028539.
 XX
 PR 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318130P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 17-SEP-2001; 2001US-0322836P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324869P.
 PR 26-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.
 PR 17-APR-2002; 2002US-0373212P.
 PR 06-SEP-2002; 2002US-00236177.
 XX
 (CURA-) CURAGEN CORP.
 XX
 PA
 XX
 PI Syytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
 PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
 PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
 PI Rieger DK, Taupier RJ, Shency SG, Liu X, Padigaru M, Alsobrook JP;
 PI Lepley DM, Edinger SR, Burgess CE;
 XX
 WPI; 2003-313242/30.
 DR P-PSDB; ABO14646.
 XX
 PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOV)
 PT and polynucleotides, useful in gene therapy, e.g. for treating or
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

```
PT stroke or infections.
XX Claim 20; Page 125-126; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful as therapeutics,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This sequence encodes a novel human NOV protein
XX
SQ Sequence 628 BP; 189 A; 157 C; 147 G; 135 T; 0 U; 0 Other;
Query Match 22.6%; Score 342; DB 10; Length 628;
Best Local Similarity 99.7%; Pred. No. 1.9e-146;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 134 GATACCTACAAATTACCGCACAGATTAATAGAAAAGAAAGAGACCGCAATTAAATGAA 193
Db 133 GATACCTACAAATTACCGCACAGATTAATAGAAAAGAAAGAGACCGCAATTAAATGAA 192
QY 194 TCCATTGCTCAGCTGAAAGATTACTGCTGACATCTGAATTTACAACTCTGGACAT 253
Db 193 TGCATTGCTCAGCTGAAAGATTACTGCTGACATCTGAATTTACAACTCTGGACAT 252
QY 254 CTGGAGAAAGCTGTAGTCTTGAATTAATTTGAAACACTTAAAGCTTTTAAACGCCCTTA 313
Db 253 CTGGAGAAAGCTGTAGTCTTGAATTAATTTGAAACACTTAAAGCTTTTAAACGCCCTTA 312
QY 314 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGGACGATCTCTGAAATCG 373
Db 313 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGGACGATCTCTGAAATCG 372
QY 374 CCCATTTCAGTCCGACTTGGATCGGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 433
Db 373 CCCATTTCAGTCCGACTTGGATCGGTTCCACTCGGATTTCAAACATGCGCCAAAGAGTC 432
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCCGGTGTGTCAGCTG 493
Db 433 TTGCAATACCTCTCCCGGTTTGAGGGCTGGACACCCAGGGAGCCCGGTGTGTCAGCTG 492
QY 494 ATCAACCACTTGCAGCGCTGGCCACCCAGTTC 526
Db 493 ATCAACCACTTGCAGCGCTGGCCACCCAGTTC 525
RESULT 15
ID AAS24787
XX AAS24787 standard; cDNA; 285 BP.
XX
AC AAS24787;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human ovarian PCR-subtracted cDNA library clone #968.
XX
XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW Gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
```

```
XX Homo sapiens.
OS WO200157207-A2.
PN 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US003733.
XX
XX 04-FEB-2000; 2000US-0180403P.
PR 28-MAR-2000; 2000US-0192745P.
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Mannion J;
XX WPI; 2001-488879/53.
XX
XX New polynucleotides encoding ovarian tumor proteins, useful for treating
XX ovarian cancer, and as probes, primers, and markers of cancer
XX progression.
XX
XX Example 1; Page 264; 378pp; English.
XX
CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumor
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumor polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumor
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumor DNA or protein by incubating isolated T-
CC cells allowing them to proliferate, and administering to the patient. The
CC sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumor protein cDNA clones
XX
SQ Sequence 285 BP; 82 A; 68 C; 66 G; 68 T; 0 U; 1 Other;
Query Match 18.1%; Score 274; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 3e-115;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 ACTGCTTGAACATCTGAAATTTGACAACTCTGGACATCTGGAGAAAGCTGTAGTCTTGA 276
Db 12 ACTGCTTGAACATCTGAAATTTGACAACTCTGGACATCTGGAGAAAGCTGTAGTCTTGA 71
QY 277 ATTAACCTTTGAAACACTTAAAGCTTTTAAACCGCTTTAACCGAGCAACAGCATCAGAAGAT 336
Db 72 ATTAACCTTTGAAACACTTAAAGCTTTTAAACCGCTTTAACCGAGCAACAGCATCAGAAGAT 131
QY 337 AATTGCTTTACAGAAATGGGAGCGATCTCTGAATCGCCATTCAGTCGACCTTGGATGC 396
Db 132 AATTGCTTTACAGAAATGGGAGCGATCTCTGAATCGCCATTCAGTCGACCTTGGATGC 191
QY 397 GTTCCACTCGGGATTTCAAAACATGCGCCAAAGAAAGCTTTGCAATACCTCTCCCGGTTGA 456
Db 192 GTTCCACTCGGGATTTCAAAACATGCGCCAAAGAAAGCTTTGCAATACCTCTCCCGGTTGA 251
QY 457 GAGCTGACACCCAGGAGCGCGGTGTGTCAG 490
Db 252 GAGCTGACACCCAGGAGCGCGGTGTGTCAG 285
```

Search completed: December 17, 2004, 05:53:30
Job time : 693.631 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:15:04 ; Search time 6271.88 Seconds
(without alignments)
11392.870 Million cell updates/sec

Title: US-10-078-650-11

Perfect score: 1511

Sequence: 1 catggacgaaggaattctc.....gtgagaagttaataaccct 1511

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sv.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1511	100.0	1511	6	BD011867
2	1378	91.2	3641	6	BD011858 Novel bHL
3	1378	91.2	3641	9	AB044088
4	1327	87.8	3431	9	BC025968 Homo sapi
5	1323	87.6	2304	6	AX591648 Sequence
6	1159	76.7	204228	9	AC022509 Homo sapi
7	725	48.0	3274	6	BD176928 Novel bHL
8	552	36.5	909	6	CQ727066 Sequence
9	274	18.1	285	6	AX209128 Sequence
10	274	18.1	397	6	AX198590 Sequence
11	143	9.5	3455	4	AY204568 Canis fam
12	128	8.5	260	6	CQ687160 Sequence
13	128	8.5	260	6	CQ688754 Sequence
14	128	8.5	260	6	CQ695947 Sequence
15	128	8.5	261	6	CQ687963 Sequence
16	128	8.5	265	6	CQ886173 Sequence
17	128	8.5	278	6	CQ895845 Sequence
18	128	8.5	502	6	AX892348 Sequence
19	128	8.5	502	6	BD027881 Sequence

20	113	7.5	5245	4	AY204567	AY204567 Canis fam
21	77	5.1	311	6	CQ696097	CQ696097 Sequence
22	54	3.6	692	10	BC037742	BC037742 Mus muscu
23	54	3.6	1421	6	BD011868	BD011868 Novel bHL
24	54	3.6	1421	10	AB044090	AB044090 Mus muscu
25	46	3.6	3101	10	AF009329	AF009329 Rattus no
26	46	3.0	300	6	BD220060	BD220060 Human gen
27	46	3.0	13281	10	AB126167	AB126167 Mus muscu
28	46	3.0	212716	2	AC144936	AC144936 Mus muscu
29	46	3.0	272078	2	AC138117	AC138117 Mus muscu
30	44	2.9	236385	2	AC110349	AC110349 Rattus no
31	33	2.2	439	10	MMBMPG1	AF100904 Mus muscu
32	33	2.2	439	10	MMBMPG1	AF028335 Mus muscu
33	33	2.2	664	10	MMU131946	AF028335 Mus muscu
34	33	2.2	671	4	AY677108	AY677108 Dasypus s
35	33	2.2	161757	10	AC128700	AC128700 Mus muscu
36	33	2.2	204856	2	AC119275	AC119275 Mus muscu
37	33	2.2	211163	2	AC141508	AC141508 Rattus no
38	32	2.1	160	9	S70397	S70397 [FRA16A, fo
39	32	2.1	390	4	AF083096	AF083096 Canis fam
40	32	2.1	393	4	AF083095	AF083095 Felis cat
41	32	2.1	3170	8	AK066310	AK066310 Oryza sat
42	32	2.1	8147	6	AX203334	AX203334 Sequence
43	32	2.1	8147	6	AX777237	AX777237 Sequence
44	32	2.1	8147	9	HSU85962	HSU85962 Homo sapien
45	32	2.1	35687	9	AC004779	AC004779 Homo sapi

ALIGNMENTS

RESULT 1
BD011867
LOCUS BD011867 1511 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011867
VERSION BD011867.1 GI:22092056
KEYWORDS Novel bHLH type transcriptional gene, DEC2
SOURCE WO 0114551-A/10.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1511)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 10 01-MAR-2001;
CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/10
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99P 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1511
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 1511; DB 6; Length 1511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGAGGAGGAGTTCCTCATTTGCAAGGAGACAGTCTACTGGAACATAGATTAT 60

Db 1 CATGGAGGAGGAGTTCCTCATTTGCAAGGAGACAGTCTACTGGAACATAGATTAT 60

Qy 61 AGGACTGGACTATTCTCTTTGTATATGTATAACCCAAAGGAGCATGAACGACGAC 120

Db 61 AGAGCTGGAGCTATCTCTTTGTATGTATGTAACCCAAAGGAGCATGAAACGAGACA 120
QY 121 CACCAAGGTAGTATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGA 180
Db 121 CACCAAGGTAGTATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGA 180
QY 181 CCGAATTAATGATGATGCTCAGCTGAAGATTTACTGCTGAACATCTGAAATGAC 240
Db 181 CCGAATTAATGATGATGCTCAGCTGAAGATTTACTGCTGAACATCTGAAATGAC 240
QY 241 AACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTTGAAACACTTAAAGC 300
Db 241 AACTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAATTAACCTTTGAAACACTTAAAGC 300
QY 301 TTTAACCGCTTTAACCGAGCAACAGATCAGAAAGATTAATGCTTTACAGATGGGAGC 360
Db 301 TTTAACCGCTTTAACCGAGCAACAGATCAGAAAGATTAATGCTTTACAGATGGGAGC 360
QY 361 ATCTCTGAATCGCCATTCAGTCCGACTTGATCGCTTCCACTCGGATTTCAAACATG 420
Db 361 ATCTCTGAATCGCCATTCAGTCCGACTTGATCGCTTCCACTCGGATTTCAAACATG 420
QY 421 CGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCGAGGAGCGC 480
Db 421 CGCCAAAGAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCGAGGAGCGC 480
QY 481 GTGTGTCCAGCTATCAACACTTGCACGCTGCGGACCCAGATTTGCCCACCCGCA 540
Db 481 GTGTGTCCAGCTATCAACACTTGCACGCTGCGGACCCAGATTTGCCCACCCGCA 540
QY 541 GGTGTTGACTCAACAGTCCCTCTGAGCAAGGACCGCGCTCCCTCGCGCGCGGTC 600
Db 541 GGTGTTGACTCAACAGTCCCTCTGAGCAAGGACCGCGCTCCCTCGCGCGCGGTC 600
QY 601 CGCGCGCGCTGCTGAGCGCGCGGAGAGAGTGGAGCCCTCGCTACTGCTGCT 660
Db 601 CGCGCGCGCTGCTGAGCGCGCGGAGAGTGGAGCCCTCGCTACTGCTGCT 660
QY 661 GCGCGTCACTCAGCGGACTCAGCCAGCCGCGGAGTCCGCGGAGAGAGAGAGAGAG 720
Db 661 GCGCGTCACTCAGCGGACTCAGCCAGCCGCGGAGTCCGCGGAGAGAGAGAGAGAG 720
QY 721 CGACAGCGCTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 CGACAGCGCTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 GGGGGGAGCGCGCTCAACATCAAGCAGAGAGCTCCGCGGAGAGAGAGAGAGAGAG 840
Db 781 GGGGGGAGCGCGCTCAACATCAAGCAGAGAGCTCCGCGGAGAGAGAGAGAGAGAG 840
QY 841 GAGGATGAAGTGAATTCCTGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 GAGGATGAAGTGAATTCCTGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 GCGCGGAG 960
Db 901 GCGCGGAG 960
QY 961 CGACGCGCGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 CGACGCGCGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 GAGCG 1080
Db 1021 GAGCG 1080
QY 1081 AGCTGCGCGCTAGCTGAGCGCTTCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AGCTGCGCGCTAGCTGAGCGCTTCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GCGCGGCTGCG 1200

Db 1141 GCGCGGCTGCG 1200
QY 1201 GCGCGGAGCG 1260
Db 1201 GCGCGGAGCG 1260
QY 1261 GTTGTGCGCGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 GTTGTGCGCGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GCGCGCGCTTGGGCG 1380
Db 1321 GCGCGCGCTTGGGCG 1380
QY 1381 CG 1440
Db 1381 CG 1440
QY 1441 AAAGGAGTCTTCCGTAATCTTGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 AAAGGAGTCTTCCGTAATCTTGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TAAATACCT 1511
Db 1501 TAAATACCT 1511

RESULT 2
LOCUS BD011858 3641 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel bHLH type transcriptional gene, DEC2.
ACCESSION BD011858
VERSION BD011858.1 GI:22092047
KEYWORDS WO 0114551-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3641)
AUTHORS Fujimoto, K., Shin, M. and Kato, Y.
TITLE Novel bHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 1 01-MAR-2001;
CHUGAI PHARM CO LTD, KATSUMI FUJIMOTO, MEI SHIN, YUKIO KATO
COMMENT OS Homo sapiens (human)
PN WO 0114551-A/1
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99P 233286
PI KATSUMI FUJIMOTO, MEI SHIN, YUKIO KATO
PC C12N15/12, C07L14/475, C07K16/18
CC

FEATURES
FT CDS Location/Qualifiers
1..3641
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.2%; Score 1378; DB 6; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGAGAGAGAGAGAGAGAG 193
Db 261 GATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGAGAGAGAGAGAGAG 320
QY 194 TGCATTGCTCAGCTGAAAGATTTTACTGCTGAACTCTGAAATTTGCAACTCTGGACAT 253
Db 321 TGCATTGCTCAGCTGAAAGATTTTACTGCTGAACTCTGAAATTTGCAACTCTGGACAT 380
QY 254 CTGAGAAAGCTGATGCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTAACCCTTA 313

Db	381	CTGGAGAAAGCTGTGTAGTCTTGGAAATTAACTTTGAACAACCTTAAAGACTTTAAACGCCCTTA	444
Qy	314	ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATTGGGAGAGCATCTCTGAAATCG	373
Db	441	ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATTGGGAGAGCATCTCTGAAATCG	500
Qy	374	CCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGAATTCAACAATGCGCCAAAGAATC	433
Db	501	CCCATTCAGTCCGACTTGGATGCGTTCCACTCGGGAATTCAACAATGCGCCAAAGAATC	560
Qy	434	TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG	493
Db	561	TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCCGGTGTGTCCAGCTG	620
Qy	494	ATCAACACATTTGCAGCGCGTGGCGACCCAGTTCTTTGCCACCCCGCAGCTGTGTACTCAA	553
Db	621	ATCAACACATTTGCAGCGCGTGGCGACCCAGTTCTTTGCCACCCCGCAGCTGTGTACTCAA	680
Qy	554	CAGGTCCCTCTGAGCAAAAGGCACCGCGCTCCCTCGGCCGCGGCTCCGCGCCGCCCC	613
Db	681	CAGGTCCCTCTGAGCAAAAGGCACCGCGCTCCCTCGGCCGCGGCTCCGCGCCGCCCC	740
Qy	614	TGCTTGAGAGCGCGGGGCGAAGATGAGCCCTTCGCTACTGCGTCCCGTATCCAG	673
Db	741	TGCTTGAGAGCGCGGGGCGAAGATGAGCCCTTCGCTACTGCGTCCCGTATCCAG	800
Qy	674	CGGACTCAGCCACGCGCGAGCTCGCGCGGAGACGACACGGACACCGACGCGGCTAC	733
Db	801	CGGACTCAGCCACGCGCGAGCTCGCGCGGAGACGACACGGACACCGACGCGGCTAC	860
Qy	734	GGCGCGAAGCCGAGGCCCGCCGACCGCAGAAAGCAAGGCGCGGGGCGAGCGCG	793
Db	861	GGCGCGAAGCCGAGGCCCGCCGACCGCAGAAAGCAAGGCGCGGGGCGAGCGCG	920
Qy	794	GTCAACCATCAGCAGGAGCTCTCCGGGAGAGTCTCGCGGCGCCCAAGAGATGAAGTG	853
Db	921	GTCAACCATCAGCAGGAGCTCTCCGGGAGAGTCTCGCGGCGCCCAAGAGATGAAGTG	980
Qy	854	GATTCCCGCGCGCGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	913
Db	981	GATTCCCGCGCGCGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1040
Qy	914	GCGCTTCTGGGGCCCGACCTTGCCTGCGCGCGCGCGCTGTGAGACCGACCGCGCCTG	973
Db	1041	GCGCTTCTGGGGCCCGACCTTGCCTGCGCGCGCGCGCTGTGAGACCGACCGCGCCTG	1100
Qy	974	CTCAGCTCGCTGTGTGGGTTTCGGCGGAGCGGAGCGCGCGCTTCCCGCAGCGCGCG	1033
Db	1101	CTCAGCTCGCTGTGTGGGTTTCGGCGGAGCGGAGCGCGCGCTTCCCGCAGCGCGCG	1160
Qy	1034	GCGCGGCCCTTCTGCCCTTCTGCTTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTG	1093
Db	1161	GCGCGGCCCTTCTGCCCTTCTGCTTCTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTG	1220
Qy	1094	GTGCAGGCCCTCTGTGACAGAGCGGCTTGAGAGATCTGTATCCCGCGCGCGGCTGCC	1153
Db	1221	GTGCAGGCCCTCTGTGACAGAGCGGCTTGAGAGATCTGTATCCCGCGCGCGGCTGCC	1280
Qy	1154	GCCCGCTTCCGCTGCTATACCCCGGATCCCGCGCGCGCGCGCGCGCGCGCGCGCG	1213
Db	1281	GCCCGCTTCCGCTGCTATACCCCGGATCCCGCGCGCGCGCGCGCGCGCGCGCGCG	1340
Qy	1214	GCGCGCGCTGCGCGCGCGCGCGCGGTTCCCTTGCCTTCTGCTGCTGCTGCTGCTGCTG	1273
Db	1341	GCGCGCGCTGCGCGCGCGCGCGCGGTTCCCTTGCCTTCTGCTGCTGCTGCTGCTGCTG	1400
Qy	1274	CCCGAGAGCGCGCGCGCGCGCGGACCTCTGCGGACGAGGTTGGCGCCCTTGGG	1333
Db	1401	CCCGAGAGCGCGCGCGCGCGGACCTCTGCGGACGAGGTTGGCGCCCTTGGG	1460
Qy	1334	GCGCGCACCCCGACGACCGCGCGGACCCACTTGCCCTTCGCGCGCGCGCGCGGAG	1393

Db 1461 GCGCGCACCCCCAGCACCCCGACGGCCGACCCACCTGCGCCGGCCCCCGCGAG 1520

QY 1394 CCGGGGAAACCGGAGAGCTCTGCTCAGAAAGATCCCTCGCAGCAGGAGAAAGAGTCTCC 1453

Db 1521 CCGGGGAAACCGGAGAGCTCTGCTCAGAAAGATCCCTCGCAGCAGGAGAAAGAGTCTCC 1580

QY 1454 TGAATCTCTGCGTCCCGAAGACGAGAGGTTCAAGCAGAGTGAGAGTTAAAAATACCT 1511

Db 1581 TGAATCTCTGCGTCCCGAAGACGAGAGGTTCAAGCAGAGTGAGAGTTAAAAATACCT 1638

RESULT 3

AB044088

LOCUS AB044088 3641 bp mRNA linear PRI 27-JAN-2001

DEFINITION Homo sapiens mRNA for bHLH protein DEC2, complete cds.

ACCESSION AB044088

VERSION AB044088.1 GI:12583668

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (sites) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Fujimoto,K., Shen,M., Noshito,M., Matsubara,K., Shingu,S., Honda,K., Yoshida,E., Suardita,K., Mateuda,Y. and Kato,Y.

TITLE Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix proteins

JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 164-171 (2001)

MEDLINE 21092582

PUBMED 11162494

REFERENCE 2 (bases 1 to 3641) Fujimoto,K.

AUTHORS Direct Submission

TITLE Submitted (31-MAY-2000) Katsumi Fujimoto, Hiroshima University, department of Biochemistry, school of dentistry; 1-2-3, Kasumi, Minami-Ku, Hiroshima 734-8553, Japan

JOURNAL (E-mail:k.fujimoto@hiroshima-u.ac.jp, Tel:81822575629, Fax:81822575629)

FEATURES

source

1..3641 Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1..3641

/gene="DEC2"

135..1583

/gene="DEC2"

/codon_start=1

/product="bHLH protein DEC2"

/protein_id="BAB21502.1"

/db_xref="GI:12583669"

/translation="MDEGPHLQEROLLEHDFIGLDYSLYMCKPKRSMKRDYDTKT YKLPHLEIKK3RDRINECIAOLKOLLPEHLKLTTLGHLEKAVILETLKHLKALTAL TQEQKHIALONGERSLSPQTSOLDLADHSGFQTCARVKQLVYLSFESWTPREPCRV QLNIHLHVAATQFLPTPQLLTQQVPLUSKGTGAPSAAGSAAPCLERAGAGOKLBPAYCV FVKIQTQPSAEAAENDTDDSGYGAEARDPREKKGAGASRVTIKQEPGEDGPA PKRMKLDPSGGSGGGGAAAAAALGPDPAALALRPDAALLSSLVAFGGGGP APFQPPAAAAAFLCLPFCFLSPSAAAAVQPFELDKSGLEKLYPAAAAAPFLLLYPGI PAPAAAAAATAAAAAAFAAPCLSSVLSPPEPKAGAAAAATLLPHEVAFLGAPHPQHPH GRTHLPFAGPRFPGNPESSAQEDPSQPGKEAP"

ORIGIN

Query Match 91.2%; Score 1378; DB 9; Length 3641;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTCAAAATTACCGCACAGATTAAAGAAAGAAAGAGACACCAATTATGA 193

Db 261 GATACCTCAAAATTACCGCACAGATTAAAGAAAGAAAGAGACACCAATTATGA 320

QY 194 TGCATTGCTCAGCTGAAGATTACTGCCTGAACATCTGAAATTGCAACTCTGGGACAT 253

Db 321 TGCATTGCTCAGCTGAAGATTACTGCCTGAACATCTGAAATTGCAACTCTGGGACAT 380

QY	254	CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAACACTTAAAGCTTTAAACGCGCTTA	313	1334	GGCCGACACCCCGACGACCCGACGCGCCACCCACCTGCGCTTCCGCGGCGCCGCGAG	1393
Db	381	CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAACACTTAAAGCTTTAAACGCGCTTA	440	1461	GGCCGACACCCCGACGACCCGACGCGCCACCCACCTGCGCTTCCGCGGCGCCGCGAG	1520
QY	314	ACCGAGCAACAGCATCAGAGATAATTTGCTTTACAGAAATGGGAGCGATCTCTGAATTCG	373	1394	CGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGACGCGAGGAAGAGCTCC	1453
Db	441	ACCGAGCAACAGCATCAGAGATAATTTGCTTTACAGAAATGGGAGCGATCTCTGAATTCG	500	1521	CGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGACGCGAGGAAGAGCTCC	1580
QY	374	CCCATTCAGTCGACCTTGATCGCTTCACTCGGATTTCAACATGCGCCAAAGAGTC	433	1454	TGAATCTTTCGCTCCCGAAGGAGCGAGGTTCAAGCAGAGTGAAGTTAAATAACCT	1511
Db	501	CCCATTCAGTCGACCTTGATCGCTTCACTCGGATTTCAACATGCGCCAAAGAGTC	560	1581	TGAATCTTTCGCTCCCGAAGGAGCGAGGTTCAAGCAGAGTGAAGTTAAATAACCT	1638
QY	434	TTGCAATACCTCTCCCGGTTTCAGAGCTGGACACCCAGGAGCGCGGTGTGTCCAGCTG	493	RESULT 4		
Db	561	TTGCAATACCTCTCCCGGTTTCAGAGCTGGACACCCAGGAGCGCGGTGTGTCCAGCTG	620	BC025968		
QY	494	ATCAACCACTTGACAGCGCTGCGCACCCAGTTCTTTGCCACCCCGCAGCTGTGTACTCAA	553	LOCUS		
Db	621	ATCAACCACTTGACAGCGCTGCGCACCCAGTTCTTTGCCACCCCGCAGCTGTGTACTCAA	680	DEFINITION		
QY	554	CAGGTCCTCTGAGCAAGCGACCGCGCTCCTCGCGCGCGGTTCGCGCGCGCGCC	613	ACCESSION		
Db	681	CAGGTCCTCTGAGCAAGCGACCGCGCTCCTCGCGCGCGGTTCGCGCGCGCGCC	740	VERSION		
QY	614	TGCTTGAGCGCGCGGCGAGAGCTTGGAGCGCCCTCGCTTCTGCTGCTCCGTCATCCAG	673	KEYWORDS		
Db	741	TGCTTGAGCGCGCGGCGAGAGCTTGGAGCGCCCTCGCTTCTGCTGCTCCGTCATCCAG	800	SOURCE		
QY	674	CGGACTCAGCGCGCGCGAGCTCGCGCGCGAGAGCAGCAGCAGCAGCGGCTAC	733	ORGANISM		
Db	801	CGGACTCAGCGCGCGCGAGCTCGCGCGCGAGAGCAGCAGCAGCAGCGGCTAC	860			
QY	734	GGCGCGAAGCGCGCGCGCGAGCGCGCGCGAGAGGCAAGGCGCGCGCGCGCGCG	793	REFERENCE		
Db	861	GGCGCGAAGCGCGCGCGCGAGCGCGCGCGAGAGGCAAGGCGCGCGCGCGCGCG	920	AUTHORS		
QY	794	GTCAACATCAAGCAGAGCGCTTCCGCGGAGAGCTCGCGCGCGCGCGCGCGCGCG	853			
Db	921	GTCAACATCAAGCAGAGCGCTTCCGCGGAGAGCTCGCGCGCGCGCGCGCGCGCG	980			
QY	854	GATTCG	913			
Db	981	GATTCG	1040			
QY	914	GGCTTTCTGGGCG	973	TITLE		
Db	1041	GGCTTTCTGGGCG	1100	JOURNAL		
QY	974	CTCAGCTCGCTGTGCGGTTCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1033	PUBMED		
Db	1101	CTCAGCTCGCTGTGCGGTTCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1160	REFERENCE		
QY	1034	GGCG	1093	AUTHORS		
Db	1161	GGCG	1220	TITLE		
QY	1094	GTGAGCGCTTCTTGACAAAGCGCGCTGAGAGATATCTGACCGCGCGCGCGCTGC	1153	JOURNAL		
Db	1221	GTGAGCGCTTCTTGACAAAGCGCGCTGAGAGATATCTGACCGCGCGCGCGCTGC	1280			
QY	1154	GGCGCGCTTCTTGATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1213			
Db	1281	GGCGCGCTTCTTGATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1340			
QY	1214	GGCGCGCTTCTTGACAAAGCGCGCTGAGAGATATCTGACCGCGCGCGCGCTGC	1273			
Db	1341	GGCGCGCTTCTTGACAAAGCGCGCTGAGAGATATCTGACCGCGCGCGCGCTGC	1400			
QY	1274	CCGAGAGAGCG	1333			
Db	1401	CCGAGAGAGCG	1460			

BC025968 3431 bp mRNA linear PRI 30-JUN-2004
Homo sapiens basic helix-loop-helix domain containing, class B, 3,
mRNA (cDNA clone MGC:39365 IMAGE:4650325), complete cds.
BC025968
BC025968.1 GI:19684063
MGC
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 3431)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,K.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetzl,E., Brownstein,M.J., Usdin,T.B., Tschiyuki,S.,
Carroll,J.P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Small,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marz,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3431)
Strausberg,R.
Direct Submission
Submitted (18-VAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncbi.nih.gov>
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
Parvaneh Saeedi, JR Stott, Angeliq Schnerch, Ursula Skalska,
Duane Smalish, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 42 Row: f Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene Prediction.

FEATURES

Location/Qualifiers

1..3431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39365 IMAGE:4650325"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOT7"

Gene

1..3431
/gene="BHLHB3"
/note="synonyms: SHARP-1, DEC2, SHARP1"
/db_xref="LocusID:79365"
/db_xref="MIM:606200"
292..1740

CDS

292..1740
/gene="BHLHB3"
/codon_start=1
/product="basic helix-loop-helix domain containing, class 3, 3"
/protein_id="AAH25968.1"
/db_xref="GI:19684064"
/db_xref="LocusID:79365"
/db_xref="MIM:606200"
/translation="MDEGIPHLQERLLEHRRFDIGLYSSLYMCKPXRSMKRDPTKDT
YKLPRLIEKRRDRINECTAOLKDLPEHLKLTTLGHLEKAVLETLKHLKALTAL
TQQHQKIITALONGERSLSPISODLAPHSFOTCAKEVLYLSRPSWTPRPRCV
QINLHHAVALFTLPQLTQVFLSGTGAPSAAGAAAFCLERACKLEPLAYCV
PIVQRTQSAELAAENDTDSYGEABARDREKKGAGASRTIIOBPFGEDSPA
PKRMKLDSSGGSGGGGAVAAAALIGPDAAAALLPDAALLSLVAFGGGG
APFPDAAAAPFCLPFLSPSAAAYVQPLDKSGLEKLYPAAAAPFLPYPGI
PAPAAAAAATAAAAAAFPCLSVSPPEKAGAAAATLLPHEVAFILGAPHPQHPH
GRTHLPFAGPREPNPESSAQEDPSQPKPEP"

ORIGIN

Query Match 87.8%; Score 1327; DB 9; Length 3431;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1377; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	134	GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACACCGAATTAATGAA	193
DB	418	GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACACCGAATTAATGAA	477
QY	194	TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGACAT	253
DB	478	TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGACAT	537
QY	254	CTGGAGAAAGCTGTAGCTCTGGATTAACCTTTGAACACTTAAAGCTTTAACCCCTTA	313
DB	538	CTGGAGAAAGCTGTAGCTCTGGATTAACCTTTGAACACTTAAAGCTTTAACCCCTTA	597
QY	314	ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGATGGGAGCGATCTCTGAATCG	373
DB	598	ACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGATGGGAGCGATCTCTGAATCG	657
QY	374	CCCATTCAGTCCGACTTGAGATGCGTTCGATCGGATTTCAAACATGCGCCAAAAGATC	433
DB	658	CCCATTCAGTCCGACTTGAGATGCGTTCGATCGGATTTCAAACATGCGCCAAAAGATC	717
QY	434	TTGCAATACCTCTCCCGTTTGAGAGCTGAGACACCGAGGAGCCGGTGTCTCCAGCTG	493
DB	718	TTGCAATACCTCTCCCGTTTGAGAGCTGAGACACCGAGGAGCCGGTGTCTCCAGCTG	777
QY	494	ATCAACCACTTGACACCGCTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGTGACTCAA	553
DB	778	ATCAACCACTTGACACCGCTGGCCACCCAGTTCTTGCCACCCCGCAGCTGTGTGACTCAA	837

RESULT 5

AX591648

LOCUS

DEFINITION

AX591648

ACCESSION

VERSION

AX591648.1

KEYWORDS

SOURCE

Sequence 9 from Patent WO0246409.

2304 bp

DNA

linear

PAT 27-JAN-2003

GI:27950039

Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Guo,X., Li,L., Patturajan,M., Shimkets,R.A., Casman,S.J.,
Malynkar,U.M., Tchernev,V.T., Vernet,C.A., Spytek,K.A.,
Shenoy,S.G., Altschrock,J.P., Edinger,S., Feyman,J.A., Stone,D.J.,
Ellerman,K., Gangoli,E.A., Boldog,F.L., Colman,S.D., Eison,A.J.,
Liu,X., Padigaru,M., Spaderna,S.K. and Zerhusen,B.D.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0246409-A 9 13-JUN-2002;
Curagen Corporation (US)

FEATURES
source Location/Qualifiers
1..2304
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 87.6%; Score 1323; DB 6; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACCGAATTAATGAA 193
Db 982 GATACCTACAAATTACCGCACAGATTAAAGAAAAGAAAGAGACCGAATTAATGAA 1041

QY 194 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 253
Db 1042 TGCATTGCTCAGCTGAAAGATTACTGCCTGAACATCTGAAATTGACAACTCTGGGACAT 1101

QY 254 CTGGAGAAAGCTGPAAGTCTTGGAAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCCTTA 313
Db 1102 CTGGAGAAAGCTGPAAGTCTTGGAAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCCTTA 1161

QY 314 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 373
Db 1162 ACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAAATCG 1221

QY 374 CCATTTCAGTCGCACATTGGATGGTTCCTACCTCGGGATTTCAACATCGGCCCAAGAGTC 433
Db 1222 CCATTTCAGTCGCACATTGGATGGTTCCTACCTCGGGATTTCAACATCGGCCCAAGAGTC 1281

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTG 493
Db 1282 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGTGTCCAGCTG 1341

QY 494 ATCAACCACTTGACAGCGCTGGCCACCAGTCTTGCCCAACCCGAGAGCTGTTGACTCAA 553
Db 1342 ATCAACCACTTGACAGCGCTGGCCACCAGTCTTGCCCAACCCGAGAGCTGTTGACTCAA 1401

QY 554 CAGTCCCTCTGAGCAAAAGCACCGGGCTCCCTCGGCGCGGGTCCGGCGCGCGCCCC 613
Db 1402 CAGTCCCTCTGAGCAAAAGCACCGGGCTCCCTCGGCGCGGGTCCGGCGCGCGCCCC 1461

QY 614 TGCTTGAGCGCGCGGGGCAAGAGCTGGAGCCCTCTGCCTACTGCTGTGCCGTATCCAG 673
Db 1462 TGCTTGAGCGCGCGGGGCAAGAGCTGGAGCCCTCTGCCTACTGCTGTGCCGTATCCAG 1521

QY 674 CGGACTCAGCCACGCGCGAGCTCGCGCGAGAACGACACGACACCGACACGGCGGTAC 733
Db 1522 CGGACTCAGCCACGCGCGAGCTCGCGCGAGAACGACACGACACCGACACGGCGGTAC 1581

QY 734 GGGCGGGAAGCCGAGGCCCGGCGGACCGGAGAAAGGCAAGCGCGGGGGGAGCGCG 793
Db 1582 GGGCGGGAAGCCGAGGCCCGGCGGACCGGAGAAAGGCAAGCGCGGGGGGAGCGCG 1641

QY 794 GTCAACATCAAGCAGAGACCTCCGGGGAGAGCTCGCGCGGCCCAAGACGATGAAGCTG 853
Db 1642 GTCAACATCAAGCAGAGACCTCCGGGGAGAGCTCGCGCGGCCCAAGACGATGAAGCTG 1701

QY 854 GATTCCCGCGCGCGGACG 913

Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, J., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsegue, H., Lozada, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, B., Maesey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Minor, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Mosser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwunigbo, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Picken, R., Primus, B., Pull, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshitari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Kucheralapati, R. and Gibbs, R.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

2 (bases 1 to 204228)
Worley, K.C.

REFERENCE AUTHORS TITLE JOURNAL

Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204228)
Worley, K.C.

REFERENCE AUTHORS TITLE JOURNAL

Submitted (30-NOV-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 30, 2000 this sequence version replaced gi:10945701.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found

at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Option Info -----

Version: 1.01
\$Id: bcm-qual-stats.cc,v 1.6 1998/10/27 00:34:03 kdurbin Exp
kdurbin \$

Options Specified: xfo
XGAPFile: ./xgap/hadw.final
OutFile: ./xgap/HADW.fa
FA File: ./xgap/HADW.fa
Sequence start: 10
Sequence end: 204237
Optimistic: 0
Use no qual cols: 0

----- Summary Statistics -----

Contig length: 204228
Phrap values in estimate: 203113
Average error rate (BCM-Phrap estimate): 0.000109636
Fraction of Phrap values less than 40 : 0.0179674
Number of consensus changing edits: 37
Number of N's in consensus : 0
Clip start: 11
Clip end: 204238
Num reads in contig: 2905
Num reads missing quality: 174
Num reads invalid quality: 0

Quality	Position	Surrounding Sequence
4	6263	ttttttttt(t)gagacaggt
7	89744	atttaataac(a)agggcagctg
7	89830	cttggagttt(g)agaccagctt
7	156088	gacaatcaga(a)atgaacttaa
8	6188	aactcatcag(c)taagaaaaa
8	7146	gtcagtaaac(a)agatagcat
8	89632	gtcagtaaac(a)agatagcat
8	89633	tcaggaataa(g)ccatctgtta
8	89634	caggaaatg(c)catctgttac
8	89745	atttaataac(a)ggcagctgg
8	89791	cagcattttg(g)agggctgagg
8	89792	agcatttttg(g)agggctgagg
8	89795	attttgggag(g)ctgaggtggg
8	89829	tctgtagtt(t)gagaccagcc
8	140166	ggggaggggg(g)gaggggggag
8	140167	ggggaggggg(g)gaggggggag
9	6160	aaaggaaaaa(t)aatcttggg
9	7030	gtggaggggt(g)aggggaagca
9	7097	gagggagggc(a)aggggaagag
9	7099	gggagggcaa(a)gggagaagt
9	7106	caaggaggaa(a)gtgccacac
9	7128	gattcatga(c)aaactcccca
9	7130	tctcatgaca(c)ccccccact
9	7135	tgacaactcc(c)ccactatcac
9	7136	gacaactccc(c)catatcac
9	7137	acaactccc(c)actatcac
9	7138	caactcccc(c)ctatcacaa
9	7139	aactcccc(c)ctatcacaa
9	7144	ccccctatc(a)caagaatagc
9	7148	actatcacaa(g)aatagcatgg
9	7149	ctatcacaa(g)aatagcatgg
9	7150	tatcacaa(g)tagcatgggg
9	89628	ccagtcagg(a)aaatgccatc
9	89768	atggtggctt(a)gagctgaat
9	89769	tggtagctta(c)agctgaatc


```
ORIGIN
Query Match 48.0%; Score 725; DB 6; Length 3274;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/db_xref="taxon:9606"

QY 134 GATACCTACAAATACCGACACAGATTAAATAGAAAAGAAAGAGACCGCAATTAATGAA 193
DB 261 GATACCTACAAATACCGACACAGATTAAATAGAAAAGAAAGAGACCGCAATTAATGAA 320
QY 194 TGCAATGCTCAGCTGAAAGATTACTGCTGAACTCTGAAATTTGACAACTCTGGACAT 253
DB 321 TGCAATGCTCAGCTGAAAGATTACTGCTGAACTCTGAAATTTGACAACTCTGGACAT 380
QY 254 CTGGAGAAAGCTGTAGTCTTTGAAATTAACCTTTGAAACACTTTAAACCGCTTA 313
DB 381 CTGGAGAAAGCTGTAGTCTTTGAAATTAACCTTTGAAACACTTTAAACCGCTTA 440
QY 314 ACCGAGCAACAGCATCAGAAAGATTAATGCTTTACAGAAATGGGAGCGATCTGAAATCG 373
DB 441 ACCGAGCAACAGCATCAGAAAGATTAATGCTTTACAGAAATGGGAGCGATCTGAAATCG 500
QY 374 CCATTCAGTCGACCTTGATGCTTCCACTCGGATTTCAAACATGCGCCAAAGATC 433
DB 501 CCATTCAGTCGACCTTGATGCTTCCACTCGGATTTCAAACATGCGCCAAAGATC 560
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGTGTCTCCAGCTG 493
DB 561 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGTGTCTCCAGCTG 620
QY 494 ATCAACCACTTGACGCCCTTGCCACCCAGTTCCTGCCACCCCGCAGCTGTTGACTCAA 553
DB 621 ATCAACCACTTGACGCCCTTGCCACCCAGTTCCTGCCACCCCGCAGCTGTTGACTCAA 680
QY 554 CAGTCCCTCTGAGCAAGACCGCGCTCCCTCGGCCCGCGGTCGCGCGCGCGCC 613
DB 681 CAGTCCCTCTGAGCAAGACCGCGCTCCCTCGGCCCGCGGTCGCGCGCGCGCC 740
QY 614 TGCTGAGCGCGCGGGGAGAGCTGGAGCCCTCGCTACTGCTGCGCTCATCCAG 673
DB 741 TGCTGAGCGCGCGGGGAGAGCTGGAGCCCTCGCTACTGCTGCGCTCATCCAG 800
QY 674 CGGACTCAGCCACGCGCGAGCTGCGCGCGAGAACGACACGAGACACCGAGCGGTAC 733
DB 801 CGGACTCAGCCACGCGCGAGCTGCGCGCGAGAACGACACGAGACACCGAGCGGTAC 860
QY 734 GCGCGGAGCGCGAGCGCGCGAGCGCGGAGGCAAGGCGCGCGCGCGCGCGCG 793
DB 861 GCGCGGAGCGCGAGCGCGCGAGCGCGGAGGCAAGGCGCGCGCGCGCGCGCG 920
QY 794 GTCAACATCAAGCAGGAGCTCCCGGGAGGAGCTCCCGCGCGCGCGCGCGCGCG 853
DB 921 GTCAACATCAAGCAGGAGCTCCCGGGAGGAGCTCCCGCGCGCGCGCGCGCGCG 980
QY 854 GATTC 858
DB 981 GATTC 985

RESULT 8
LOCUS CQ727066
DEFINITION Sequence 13000 from Patent WO02068579.
ACCESSION CQ727066
VERSION CQ727066.1 GI:42291946
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13000 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Source 1..909
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 36.5%; Score 552; DB 6; Length 909;
Best Local Similarity 100.0%; Pred. No. 6,3e-275;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTTAAAGCT 301
DB 181 ACTCTGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTTAAAGCT 240
QY 302 TTAACCCCTTAACCGAGCAACAGCATCAGAAATTAATGCTTTACAGAAATGGGAGCGA 361
DB 241 TTAACCCCTTAACCGAGCAACAGCATCAGAAATTAATGCTTTACAGAAATGGGAGCGA 300
QY 362 TCTCTGAAATCGCCCACTTCAGTCCGACTTGGATGCTTCCACTCGGATTTCAAACATGC 421
DB 301 TCTCTGAAATCGCCCACTTCAGTCCGACTTGGATGCTTCCACTCGGATTTCAAACATGC 360
QY 422 GCCAAAGAGTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGG 481
DB 361 GCCAAAGAGTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGG 420
QY 482 TGTGTCAGCTGATCAACCACTTCGACGCGCTGGCCACCCAGTTCCTGCCACCCCGCAG 541
DB 421 TGTGTCAGCTGATCAACCACTTCGACGCGCTGGCCACCCAGTTCCTGCCACCCCGCAG 480
QY 542 CTGTTGATCAACAGTTCCTCTTGAGCAAGGACACCGCGCTCCCTCGCGCGCGCGGTCC 601
DB 481 CTGTTGATCAACAGTTCCTCTTGAGCAAGGACACCGCGCTCCCTCGCGCGCGCGGTCC 540
QY 602 GCGCGCGCGCGCTGAGCGCGCGCGGAGAGCTGGAGCCCTCGCTCGCTACTCGGTG 661
DB 541 GCGCGCGCGCGCTGAGCGCGCGCGGAGAGCTGGAGCCCTCGCTCGCTACTCGGTG 600
QY 662 CCCCTCATCCAGCGAGCTCAGCCCGCGAGCTCGCCCGGAGAACGACACGAGACAC 721
DB 601 CCCCTCATCCAGCGAGCTCAGCCCGCGAGCTCGCCCGGAGAACGACACGAGACAC 660
QY 722 GACGCGCTACGCGCGGAGCGCGAGCGCGCGCGGAGAACGCGGAGAAAGGCGCGG 781
DB 661 GACGCGCTACGCGCGGAGCGCGAGCGCGCGGAGAACGCGGAGAAAGGCGCGG 720
QY 782 GGGGCGAGCGCG 793
DB 721 GGGGCGAGCGCG 732

RESULT 9
LOCUS AX209128
DEFINITION Sequence 968 from Patent WO0157207.
ACCESSION AX209128
VERSION AX209128.1 GI:15423551
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer
JOURNAL Patent: WO 0157207-A 968 09-AUG-2001;
```

```

FEATURES             CORIXA CORPORATION (US)
    source            Location/Qualifiers
    1..285
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match          18.1%; Score 274; DB 6; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.9e-130; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 0;

QY 217 ACTGCTGAACATCTGAAATTGACAACTCTGGGACATCTGGGAAAGCTGTAGTCTTGA 276
DB 12 ACTGCTGAACATCTGAAATTGACAACTCTGGGACATCTGGGAAAGCTGTAGTCTTGA 71

QY 277 ATTAACCTTGAACACATCTTAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 336
DB 72 ATTAACCTTGAACACATCTTAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 131

QY 337 AATTGCTTTACAGAAATGGGAGGAGTCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 396
DB 132 AATTGCTTTACAGAAATGGGAGGAGTCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 191

QY 397 GTTCCACTCGGGATTTCAACATGCGCCAAAGAGTCTTTGCAATACCTCTCCCGTTTGA 456
DB 192 GTTCCACTCGGGATTTCAACATGCGCCAAAGAGTCTTTGCAATACCTCTCCCGTTTGA 251

QY 457 GACGTGACACCCAGGAGCGCGGTGTCTCCAG 490
DB 252 GACGTGACACCCAGGAGCGCGGTGTCTCCAG 285

RESULT 10
LOCUS                AX198590
DEFINITION           Sequence 1045 from Patent WO0151513.
ACCESSION            AX198590
VERSION              AX198590.1 GI:15388911
KEYWORDS
SOURCE               Homo sapiens (human)
ORGANISM              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  ALgate, P.A.
  Ovarian tumor-associated sequences
  Patent: WO 0151513-A 1045 19-JUL-2001;
  CORIXA CORPORATION (US)
FEATURES
    source            Location/Qualifiers
    1..397
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match          18.1%; Score 274; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 2e-130; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 0;

QY 217 ACTGCTGAACATCTGAAATTGACAACTCTGGGACATCTGGGAAAGCTGTAGTCTTGA 276
DB 12 ACTGCTGAACATCTGAAATTGACAACTCTGGGACATCTGGGAAAGCTGTAGTCTTGA 71

QY 277 ATTAACCTTGAACACATCTTAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 336
DB 72 ATTAACCTTGAACACATCTTAAGCTTTAAACCGCTTTAACCGGACCAACAGCATCAGAAGAT 131

QY 337 AATTGCTTTACAGAAATGGGAGGAGTCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 396
DB 132 AATTGCTTTACAGAAATGGGAGGAGTCTCTGAAATCGCCCATTCAGTCCGACTTGGATGC 191

QY 397 GTTCCACTCGGGATTTCAACATGCGCCAAAGAGTCTTTGCAATACCTCTCCCGTTTGA 456

```

```

DB 192 GTTCCACTCGGGATTTCAACATGCGCCAAAGAGTCTTTGCAATACCTCTCCCGTTTGA 251

QY 457 GACGTGACACCCAGGAGCGCGGTGTCTCCAG 490
DB 252 GACGTGACACCCAGGAGCGCGGTGTCTCCAG 285

RESULT 11
LOCUS                AY204568
DEFINITION           Canis familiaris SHARPI protein (SHARPI) mRNA, complete cds.
ACCESSION            AY204568
VERSION              AY204568.1 GI:33439703
KEYWORDS
SOURCE               Canis familiaris (dog)
ORGANISM              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
  1 (bases 1 to 3455)
  Kukekova, A.V., Aguirre, G.D. and Acland, G.M.
  Cloning and characterization of canine SHARPI and its evaluation as
  a positional candidate for canine early retinal degeneration (erd)
  Gene 312, 335-343 (2003)
  2 (bases 1 to 3455)
  Kukekova, A.V., Aguirre, G.D. and Acland, G.M.
  Direct Submission
  Submitted (18-DEC-2002) Center for Canine Genetics and
  Reproduction, James A. Baker Institute for Animal Health,
  Hungerford Hill Road, Ithaca, NY 14853-6401, USA
FEATURES
    source            Location/Qualifiers
    1..3455
    /organism="Canis familiaris"
    /mol_type="mRNA"
    /db_xref="taxon:9615"
    /tissue="retina"
    1..3455
    /gene="SHARPI"
    269..1810
    /gene="SHARPI"
    /notes="BHLH transcription factor family member; DEC2;
    BHLHB3"
    /codon_start=1
    /product="SHARPI protein"
    /protein_id="AAP12465.1"
    /db_xref="GI:33439704"
    /translation="MDEGPHLQERQLLEHRDFIGLDYSSLYNCKPKRSMKDDSDKT
    YKLPHRLIEKRRDRINECIAQLKOLLPEHLKLTTLGHLEKAVLELTLLKHLKALTAL
    TEQHQHOKITLQNGERSLKSPQSDLDLDFHSGFTCAKEVLQVLSRFESWTFREQCV
    OLINHLHVAATQFLPTPOLLTQVPLSKGAGAPSAAPAGSAAPCLERAGOKLEPLA
    HCVPIQRTQPSAELEAENDTDTDSYGGGEARPDPRDKGSGAGVTTIKGEPGED
    SPAPKMKLDSRGGGGGGGGGLGGGGGLGGGGGGGGGGGGGGGGAANAALLGPDPA
    AALLRPDAALLSLVGGGAPPAFAAANAAPFCILPFIYLSPSAAAAYVQFFLDK
    GLEKLYPAAAAAPFELLYPGIPAPAAAAAANAANAAPFCLSLVSPPEKAGAA
    AATLLPHEVAPGALHPHPHPHPHPHPHGHRTHLAFAGAREPGNPGNSAQEDPSQP
    GKEAH"

ORIGIN
Query Match          9.5%; Score 143; DB 4; Length 3455;
Best Local Similarity 100.0%; Pred. No. 2.6e-62;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 CAGCATCAGAGATTAATGCTTTTACAGAAATGGGGAGGATCTCTGAAATCGCCCATTCAG 382
DB 584 CAGCATCAGAGATTAATGCTTTTACAGATGGGAGGATCTCTGAAATCGCCCATTCAG 643

QY 383 TCGGACTTGATGCGTTCCTGCTCGGATTTCAAAACATGCGCCCAAGATCTTGCATAC 442
DB 644 TCGGACTTGATGCGTTCCTGCTCGGATTTCAAAACATGCGCCCAAGATCTTGCATAC 703

QY 443 CTCTCCCGGTTTCAGAGCTGGAC 465
DB 704 CTCTCCCGGTTTCAGAGCTGGAC 726

```

```
RESULT 12
CQ687160          CQ687160          260 bp      DNA          linear      PAT 03-FEB-2004
LOCUS              Sequence 32086 from Patent WO02070737.
DEFINITION
ACCESSION          CQ687160
VERSION            CQ687160.1 GI:42217526
KEYWORDS
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS            Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE              Compositions and methods relating to osteoarthritis
JOURNAL            Patent: WO 02070737-A 32086 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source             Location/Qualifiers
1..260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match       8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 60
Db 105 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 164
QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 120
Db 165 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 224
QY 121 CACCAAGG 128
Db 225 CACCAAGG 232
RESULT 13
CQ688754          CQ688754          260 bp      DNA          linear      PAT 03-FEB-2004
LOCUS              Sequence 33680 from Patent WO02070737.
DEFINITION
ACCESSION          CQ688754
VERSION            CQ688754.1 GI:42221719
KEYWORDS
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS            Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE              Compositions and methods relating to osteoarthritis
JOURNAL            Patent: WO 02070737-A 33680 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source             Location/Qualifiers
1..260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match       8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 60
Db 115 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 174
```

```
QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 120
Db 175 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 234
QY 121 CACCAAGG 128
Db 235 CACCAAGG 242
RESULT 14
CQ695847          CQ695847          260 bp      DNA          linear      PAT 03-FEB-2004
LOCUS              Sequence 40773 from Patent WO02070737.
DEFINITION
ACCESSION          CQ695847
VERSION            CQ695847.1 GI:42243288
KEYWORDS
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS            Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE              Compositions and methods relating to osteoarthritis
JOURNAL            Patent: WO 02070737-A 40773 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source             Location/Qualifiers
1..260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match       8.5%; Score 128; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-54;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 60
Db 116 CATGGACGAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAACATAGAGATTTAT 175
QY 61 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 120
Db 176 AGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGAGAGCATGAACGAGACGA 235
QY 121 CACCAAGG 128
Db 236 CACCAAGG 243
RESULT 15
CQ687963          CQ687963          261 bp      DNA          linear      PAT 03-FEB-2004
LOCUS              Sequence 32889 from Patent WO02070737.
DEFINITION
ACCESSION          CQ687963
VERSION            CQ687963.1 GI:42219444
KEYWORDS
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS            Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE              Compositions and methods relating to osteoarthritis
JOURNAL            Patent: WO 02070737-A 32889 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source             Location/Qualifiers
1..261
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match       8.5%; Score 128; DB 6; Length 261;
```

Best Local Similarity 100.0%; Pred. No. 1.5e-54;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTTACTGGAAACATAGAGATTTTAT	60
Db	114	CATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGTTTACTGGAAACATAGAGATTTTAT	173
Qy	61	AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA	120
Db	174	AGGACTGGACTATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGAGACGA	233
Qy	121	CACCAAGG	128
Db	234	CACCAAGG	241

Search completed: December 17, 2004, 11:50:02
 Job time : 6276.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:13:29 ; Search time 1659.37 Seconds
(without alignments)
11518.321 Million cell updates/sec

Title: US-10-078-650-1
Perfect score: 3641
Sequence: 1 ctgactgaagaggagagc.....ttaaaaaaaaaaaaaaaaaa 3641

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 23Sep04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3641	100.0	3641	4	Aaf74767 Human DEC
2	3641	100.0	3641	6	Abx76400 cDNA enco
3	3641	100.0	3641	9	Acc72416 Human ova
4	3641	100.0	3641	12	Adn05013 Antipsoi
5	3152	86.6	3521	10	Adc30675 Human nov
6	1583	43.5	2304	6	Abt05457 DNA of NO
7	1471	40.4	1471	10	Acd19338 cDNA enco
8	1378	37.8	1511	4	Aaf74776 Human DEC
9	629	17.3	753	4	Aal26629 Human bre
10	552	15.2	909	10	Abx74439 Human cdn
11	516	14.2	1573	12	Acq24375 Human sof
12	496	13.6	498	4	Aas26175 Human cdn
13	496	13.6	498	8	Abx73516 Human nov
14	430	11.8	430	9	Acc72415 Human cdn
15	437	11.7	628	10	Acd19339 cDNA enco
16	391	10.7	462	4	Aal16710 Human bre
17	386	10.6	447	4	Aal25553 Human bre
18	382	10.5	527	11	Adm10886 Human ova
19	382	10.5	527	12	Adj11216 Human ova
20	392	10.5	527	12	Adm43477 Human ova
21	368	10.1	896	10	Adc32475 Human nov

22	366	10.1	454	4	AAS26590	Human cdn
23	366	10.1	454	8	ABX73931	Human nov
24	355	9.8	521	5	AaH83267	Human ova
25	354	9.7	501	4	AAS25314	Human ova
26	341	9.4	580	4	AAS24664	Human ova
27	341	9.4	580	5	AaH83289	Human ova
28	339	9.3	411	4	AAL07807	Human bre
c 29	308	8.5	553	5	ADL43969	Human ova
c 30	308	8.5	684	5	ADL37585	Human ova
c 31	308	8.5	684	5	ADL72444	Human ova
32	296	8.1	409	4	AAS25322	Human ova
33	295	8.1	502	4	AAS24569	Human ova
34	295	8.1	502	5	AaH83185	Human ova
35	280	7.7	501	4	AAS24644	Human ova
36	274	7.5	285	5	AAS24787	Human ova
c 37	274	7.5	337	5	AaH83421	Human ova
c 38	271	7.4	322	4	AAS24447	Human ova
c 39	271	7.4	424	5	AaH83050	Human ova
c 40	266	7.3	440	6	ABL38131	Human col
41	245	6.8	502	3	AAC04136	Human sec
42	242	6.6	604	5	ADL45897	Human ova
43	216	5.9	216	6	ABV87290	Human col
44	209	5.7	312	12	ADQ20370	Human sof
45	194	5.3	524	5	ADL40167	Human ova

ALIGNMENTS

RESULT 1
AAF74767
ID AAF74767 standard; cDNA; 3641 BP.
AC AAF74767;
XX
XX 17-MAY-2001 (first entry)
XX Human DEC2a encoding cDNA sequence SEQ ID NO:1.
XX DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
XX basic helix loop helix protein; cell differentiation; proliferation; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 135..1583
XX /*tag= a
XX /product= "DEC2a"
XX /note= "bHLH type transcription factor"
XX
XX WO200114551-A1.
XX
XX 01-MAR-2001.
XX
XX 19-JUN-2000; 2000WO-JF003991.
XX
XX 19-AUG-1999; 99JP-00233286.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Fujimoto K, Shin M, Kato Y;
XX
XX WPI; 2001-202935/20.
XX
XX P-PSDB; AAB70692.
XX
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
XX development of drugs for treatment of disorders of cell differentiation
XX and proliferation.
XX
XX Claim 1; Page 48-55; 83pp; Japanese.
XX
XX The present invention describes a basic helix loop helix (bHLH) type
XX transcription factor designated DEC2. DEC2 can be used as a tool in the

CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC encodes the specifically claimed human DBC2a protein, as given in the
CC present invention
XX
SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
Query Match 100.0%; Score 3641; DB 4; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCACTGAAGAGGGAGAGCGAGAGAGACTGGAGACGACAGATCCCCCAAGGTCTC 60
DB |||||
QY 1 CTGCACTGAAGAGGGAGAGCGAGAGAGACTGGAGACGACAGATCCCCCAAGGTCTC 60
DB |||||
QY 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCAAAATCGAAACAGAGAGAAACG 120
DB |||||
QY 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCAAAATCGAAACAGAGAGAAACG 120
DB |||||
QY 121 AACAGCAGTTGAACATGGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC 180
DB |||||
QY 121 AACAGCAGTTGAACATGGAAGGAATTCCTCATTTGCAAGAGAGACAGTTACTGGAAC 180
DB |||||
QY 181 ATAGAGATTTATAGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGGAGCA 240
DB |||||
QY 181 ATAGAGATTTATAGGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAGGAGCA 240
DB |||||
QY 241 TGAACAGAGACACACCAAGGATACCTACAAATACCAGCAGAGATTATAGAAAGAAAA 300
DB |||||
QY 241 TGAACAGAGACACACCAAGGATACCTACAAATACCAGCAGAGATTATAGAAAGAAAA 300
DB |||||
QY 301 GAAGAGCCGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB |||||
QY 301 GAAGAGCCGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB |||||
QY 361 AATTGACAACTCTGGAGACTCTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACT 420
DB |||||
QY 361 AATTGACAACTCTGGAGACTCTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACT 420
DB |||||
QY 421 TAAAGCTTTAACCGCTTTAACCGAGCAACAGCATCAGAGATTAATTTGAAACACT 480
DB |||||
QY 421 TAAAGCTTTAACCGCTTTAACCGAGCAACAGCATCAGAGATTAATTTGAAACACT 480
DB |||||
QY 481 GGGAGCGATCTGAAATCGCCCATTCAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB |||||
QY 481 GGGAGCGATCTGAAATCGCCCATTCAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB |||||
QY 541 AAACATGCGCCAAAGAAAGTCTTGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
DB |||||
QY 541 AAACATGCGCCAAAGAAAGTCTTGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
DB |||||
QY 601 AGCCGGGTGCTCAGCTGATCAACCACTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 660
DB |||||
QY 601 AGCCGGGTGCTCAGCTGATCAACCACTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 660
DB |||||
QY 661 CCGCGCAGCTGTTGACTCAACAGGTCCTCTGAGCAAAAGGACACCGCGCTCCCTCGGCGG 720
DB |||||
QY 661 CCGCGCAGCTGTTGACTCAACAGGTCCTCTGAGCAAAAGGACACCGCGCTCCCTCGGCGG 720
DB |||||
QY 721 CCGGGTCCGCGCGCGCCCTGCTGAGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB |||||
QY 721 CCGGGTCCGCGCGCGCCCTGCTGAGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB |||||
QY 781 ACTGCGTCCGCTCATCAGCGGATCTCAGCCCGAGCGCGGCGGCGGCGGCGGCGGCGGCG 840
DB |||||
QY 841 CCGACACGACAGCGGCTACCGCGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGG 900
DB |||||
QY 841 CCGACACGACAGCGGCTACCGCGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGG 900
DB |||||
QY 901 AAGGCGCGGGCGAGCGCGCTCACCATCAGCAGAGCGCTCCCGGGAGGAGCTCGCGG 960
DB |||||

DB 901 AAGGCGCGGGCGAGCGCGCTCACCATCAAGCAGAGAGCTCCCGGGAGAGACTCGCCGG 960
QY |||||
DB 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGGCGGCGAGCGCGCGCGCGCGCGCGCGGGGG 1020
QY |||||
DB 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGGCGGCGAGCGCGCGCGCGCGCGCGGGGG 1020
QY |||||
DB 1021 GC CGCGCGCGCGCGAGCGCGCTTCTGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCTGC 1080
DB |||||
DB 1021 GC CGCGCGCGCGCGAGCGCGCTTCTGGGGCGGCGCGCGCGCGCGCGCGCGCGCTGC 1080
QY |||||
DB 1081 TGAGACCGGAGCGCGCGCTGCTCAGCTCGCTGGTGGCTTCCGCGGAGCGGAGCGCGCG 1140
DB |||||
DB 1081 TGAGACCGGAGCGCGCGCTGCTCAGCTCGCTGGTGGCTTCCGCGGAGCGGAGCGCGCG 1140
QY |||||
DB 1141 CCTTCCCGAGCGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTCTCTCTCTCTCT 1200
DB |||||
DB 1141 CCTTCCCGAGCGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTCTCTCTCTCT 1200
QY |||||
DB 1201 CTTCTGAGCTGCGCGCTACGTGAGCCCTTCTGGAAGAAGCGGCTGGAAGAATATC 1260
DB |||||
DB 1201 CTTCTGAGCTGCGCGCTACGTGAGCCCTTCTGGAAGAAGCGGCTGGAAGAATATC 1260
QY |||||
DB 1261 TGTACCGCGCGCGCGCTGCGCGCTTCCGCTGCTATACCCCGGATCCCGCGCGCGG 1320
DB |||||
DB 1261 TGTACCGCGCGCGCGCTGCGCGCTTCCGCTGCTATACCCCGGATCCCGCGCGCGG 1320
QY |||||
DB 1321 CGGAGCGCGCGCGCGCGCTGCGCGCTTCCGCGCGCGCGCGCGCGCGCTTCTGCTGCT 1380
DB |||||
DB 1321 CGGAGCGCGCGCGCGCGCTGCGCGCTTCCGCGCGCGCGCGCGCGCGCTTCTGCTGCT 1380
QY |||||
DB 1381 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB |||||
DB 1381 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY |||||
DB 1441 ACAGAGTGGCGCGCTTGGGGCGCGCGCACCCCGAGCACCCCGCGCGCGCGCGCGCGCT 1500
DB |||||
DB 1441 ACAGAGTGGCGCGCTTGGGGCGCGCGCACCCCGAGCACCCCGCGCGCGCGCGCGCT 1500
QY |||||
DB 1501 CTTCTCGCT 1560
DB |||||
DB 1501 CTTCTCGCT 1560
QY |||||
DB 1561 AGCCAGGAAAGAGCTCCCTGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB |||||
DB 1561 AGCCAGGAAAGAGCTCCCTGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY |||||
DB 1621 GAGAGTTAAATAFACCTTTAAGAGGTTCAAGCAGAGTGAAGTTAAATAFACCTTTAAG 1680
DB |||||
DB 1621 GAGAGTTAAATAFACCTTTAAGAGGTTCAAGCAGAGTGAAGTTAAATAFACCTTTAAG 1680
QY |||||
DB 1681 GTCTTTAAGGAGGAGTGAATAGATGACACAGAGGATTAACAGAAACAAACAAACAG 1740
DB |||||
DB 1681 GTCTTTAAGGAGGAGTGAATAGATGACACAGAGGATTAACAGAAACAAACAAACAG 1740
QY |||||
DB 1741 GTGTTATGTATCATTCGAGGTTCTGTTGCTCATCCCGCACCAACCCCGCTTCAAC 1800
DB |||||
DB 1741 GTGTTATGTATCATTCGAGGTTCTGTTGCTCATCCCGCACCAACCCCGCTTCAAC 1800
QY |||||
DB 1801 CACTACATCCCTTCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB |||||
DB 1801 CACTACATCCCTTCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY |||||
DB 1861 TCTTTTTTTTTTAAATCCCAATAAATTTTGGCGCGCTTTTGGCGCGCTTTTCCATTTCTCT 1920
DB |||||
DB 1861 TCTTTTTTTTTTAAATCCCAATAAATTTTGGCGCGCTTTTGGCGCGCTTTTCCATTTCTCT 1920
QY |||||
DB 1921 TAAATTTGAAACCTTAATTCGAGAGGAGTGAAGAGGCTGTTCTGCTGCTGCTGCTGCTG 1980
DB |||||
DB 1921 TAAATTTGAAACCTTAATTCGAGAGGAGTGAAGAGGCTGTTCTGCTGCTGCTGCTGCTG 1980
QY |||||
DB 1981 GAACCCCGGGTAGGGGAAAGATGTTAACCTTTGACGCTCTTTGGAGTTGACATGGAGAC 2040
DB |||||
DB 1981 GAACCCCGGGTAGGGGAAAGATGTTAACCTTTGACGCTCTTTGGAGTTGACATGGAGAC 2040

```
QY 2041 AGCAGGTAGTGTGTTATGTAGAGCTAGTCTCTCAAGCTGCCCTGCTGTTTATAGGAGCGT 2100
Db 2041 AGCAGGTAGTGTGTTATGTAGAGCTAGTCTCTCAAGCTGCCCTGCTGTTTATAGGAGCGT 2100
QY 2101 TCACAAAACAGATTGAGGCTCTTTTATAGAAATTTACTCTTCAGTATTTCTTAATGT 2160
Db 2101 TCACAAAACAGATTGAGGCTCTTTTATAGAAATTTACTCTTCAGTATTTCTTAATGT 2160
QY 2161 TCAGCTTTCTAAAGGCATATATTTTCAAGAGTGAGGATGCGAGTTTCTCACGTTGCA 2220
Db 2161 TCAGCTTTCTAAAGGCATATATTTTCAAGAGTGAGGATGCGAGTTTCTCACGTTGCA 2220
QY 2221 ACCATTTCTCAAGTGGTGTAAATGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACAC 2280
Db 2221 ACCATTTCTCAAGTGGTGTAAATGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACAC 2280
QY 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAAGAGATGATGAGAAAGCCAGAAATCA 2340
Db 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAAGAGATGATGAGAAAGCCAGAAATCA 2340
QY 2341 TTCCTAGTACATTTGCTAACACTTTATTGAGAAATTGACCATGAATTAATGGACTCATCT 2400
Db 2341 TTCCTAGTACATTTGCTAACACTTTATTGAGAAATTGACCATGAATTAATGGACTCATCT 2400
QY 2401 TAATTTCTTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
Db 2401 TAATTTCTTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
QY 2461 TAGATAGGTATCTATACATACATCTCAAGTGCATCTATTCCTCACTCTCATTAATCCAT 2520
Db 2461 TAGATAGGTATCTATACATACATCTCAAGTGCATCTATTCCTCACTCTCATTAATCCAT 2520
QY 2521 CATGTTCTTAATTTTGTAAATCTTACTGTAAAAAAGTAGTGCATCAACTTCAAAACAAA 2580
Db 2521 CATGTTCTTAATTTTGTAAATCTTACTGTAAAAAAGTAGTGCATCAACTTCAAAACAAA 2580
QY 2581 ACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2640
Db 2581 ACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2640
QY 2641 TTCAAAGTGAACGAAGCATTTAACTGGCCAGTTTGTATGCAAAATGCTGTAAAGATAT 2700
Db 2641 TTCAAAGTGAACGAAGCATTTAACTGGCCAGTTTGTATGCAAAATGCTGTAAAGATAT 2700
QY 2701 AGAATGAAGTCTGTGAGGCTTCTCTATCTCAAGTCTATGTAATTTCTGGAGACCAAC 2760
Db 2701 AGAATGAAGTCTGTGAGGCTTCTCTATCTCAAGTCTATGTAATTTCTGGAGACCAAC 2760
QY 2761 CAGATACCAAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGCT 2820
Db 2761 CAGATACCAAGATAATCACAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGCT 2820
QY 2821 AGATATTTTGTGTTTCCCAAGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA 2880
Db 2821 AGATATTTTGTGTTTCCCAAGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTA 2880
QY 2881 AGGGGTGACACAAACAGTCTGACTATGATGAGGAAATATCTGGTCTTTTCTGAGT 2940
Db 2881 AGGGGTGACACAAACAGTCTGACTATGATGAGGAAATATCTGGTCTTTTCTGAGT 2940
QY 2941 TTGGTGCATTTGCTGCTGTTGCTACTGTTTGGCTCAACCGCTGTGTTTAAACAAAGT 3000
Db 2941 TTGGTGCATTTGCTGCTGTTGCTACTGTTTGGCTCAACCGCTGTGTTTAAACAAAGT 3000
QY 3001 TAAACTCTTAGGCTCAAGGTGGCTCTTATGATACATAGTGTGTAATATCAATCAATTAATG 3060
Db 3001 TAAACTCTTAGGCTCAAGGTGGCTCTTATGATACATAGTGTGTAATATCAATCAATTAATG 3060
QY 3061 ATGCTGACATGCTATTTTGTAGGGAGAAATATGCTGAATGATATTTTTCAGTTAAAA 3120
Db 3061 ATGCTGACATGCTATTTTGTAGGGAGAAATATGCTGAATGATATTTTTCAGTTAAAA 3120
```

```
QY 3121 TATCTTTTGGGAGGATTTGCTGAAAAGTTTGCACCTTTTGTTCACATCTTATGCTTGGA 3180
Db 3121 TATCTTTTGGGAGGATTTGCTGAAAAGTTTGCACCTTTTGTTCACATCTTATGCTTGGA 3180
QY 3181 CAAGCTTATGCTGCTTTAAATTTTAAATAATTTAAATAATTTAAATAATTTAAATAAT 3240
Db 3181 CAAGCTTATGCTGCTTTAAATTTTAAATAATTTAAATAATTTAAATAATTTAAATAAT 3240
QY 3241 TGGTTTAAAGAAAGTTTATGATGATGACGATAAATCTAGAAATTTACCTTTATATCTAGTATT 3300
Db 3241 TGGTTTAAAGAAAGTTTATGATGATGACGATAAATCTAGAAATTTACCTTTATATCTAGTATT 3300
QY 3301 TTCAGCACTCCATAAATTTCTATTACCTTAAATTTTGCACACTATTTTGTGATTTAAAAAT 3360
Db 3301 TTCAGCACTCCATAAATTTCTATTACCTTAAATTTTGCACACTATTTTGTGATTTAAAAAT 3360
QY 3361 TCCTACTAAGCAATAAACCTTTTAAATATACGATATGATTTGCTAAATTTAAATAAAGA 3420
Db 3361 TCCTACTAAGCAATAAACCTTTTAAATATACGATATGATTTGCTAAATTTAAATAAAGA 3420
QY 3421 CATAATGGATGCTCAATTTAGTATTTTAAATATCTAATACTATAGGATACAAATCATTACA 3480
Db 3421 CATAATGGATGCTCAATTTAGTATTTTAAATATCTAATACTATAGGATACAAATCATTACA 3480
QY 3481 GTTCTCAGATTTACACCTTTTGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
Db 3481 GTTCTCAGATTTACACCTTTTGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
QY 3541 GCAAGCTCCAGGCTCGGCTTTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
Db 3541 GCAAGCTCCAGGCTCGGCTTTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
QY 3601 TGCAAAAGAAAAAACCCTTACCCTTAAATAAATAAATAAATAAATAAATAAATAAATAA 3641
Db 3601 TGCAAAAGAAAAAACCCTTACCCTTAAATAAATAAATAAATAAATAAATAAATAAATAA 3641

RESULT 2
ABS76400
ID ABS76400 standard; cDNA; 3641 BP.
XX AC ABS76400;
XX AC ABS76400;
XX DT 11-DEC-2002 (first entry)
XX DE cDNA encoding human ovarian cancer marker OV4.
XX KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
XX KW central nervous system disorder; bacterial meningitis; viral meningitis;
XX KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
XX KW brain herniation; inflammation; encephalitis; testicular disorder;
XX KW nontuberculous granulomatous orchitis; connective tissue disorder;
XX KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
XX KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200271928-A2.
XX PD 19-SEP-2002.
XX PF 14-MAR-2002; 2002WO-US007826.
XX PR 14-MAR-2001; 2001US-0276025P.
XX PR 14-MAR-2001; 2001US-0276026P.
XX PR 10-AUG-2001; 2001US-0311732P.
XX PR 19-SEP-2001; 2001US-0323580P.
XX PR 26-SEP-2001; 2001US-0324967P.
XX PR 26-SEP-2001; 2001US-0325102P.
XX PR 26-SEP-2001; 2001US-0325149P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
```

PI Monahan JE, Gannavarapu M, Hoersch S, Kamathkar S, Kovatis SG;
PI Meyers RE, Morrissey MP, Olandt Fu, Sen A, Vleby PO, Mills GB;
PI East RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR P-PSDB; ABG96308.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX Disclosure; Page 186-187; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterizing cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nonbacterial granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present nucleic acid
CC sequence encodes one of the ovarian cancer markers described in the
XX invention
XX
SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;

Query Match 100.0%; Score 3641; DB 6; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCTGAAGAGGAGCGAGAGAGACTGGAGAGCAGATCCCCCAAGGCTCTC 60
DB 1 CTGCCTGAAGAGGAGCGAGAGAGACTGGAGAGCAGATCCCCCAAGGCTCTC 60

QY 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCCCAAAATCGAAACAGAGAAACG 120
DB 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCCCAAAATCGAAACAGAGAAACG 120

QY 121 AACAGAGTTGAACATGGAGAGGAAATTCCTCATTTGCAAGAGAGAGAGTACTGGAC 180
DB 121 AACAGAGTTGAACATGGAGAGGAAATTCCTCATTTGCAAGAGAGAGAGTACTGGAC 180

QY 181 ATAGAGATTTATAGAGCTGGAGTATTCCTCTTTGTATGTATGTAAACCAAAAGGAGCA 240
DB 181 ATAGAGATTTATAGAGCTGGAGTATTCCTCTTTGTATGTATGTAAACCAAAAGGAGCA 240

QY 241 TGAACGAGAGAGACACCAAGGATACCTACAAATTAACCGACAGATTAATAGAAAGAAAA 300
DB 241 TGAACGAGAGAGACACCAAGGATACCTACAAATTAACCGACAGATTAATAGAAAGAAAA 300

QY 301 GAAGAGACCGAATTAATGAATGATTCAGTTCAGTGAAGATTTACTGCTGAACATCTGA 360
DB 301 GAAGAGACCGAATTAATGAATGATTCAGTTCAGTGAAGATTTACTGCTGAACATCTGA 360

QY 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAATCTTTGAAACACT 420
DB 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAATCTTTGAAACACT 420

DB 361 AATTGACAACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAATCTTTGAAACACT 420
QY 421 TAAAGGCTTTAAACCGGCTTAAACCGAGCAACAGATCAGAAAGATAATTTGCTTTACAGAAATG 480
DB 421 TAAAGGCTTTAAACCGGCTTAAACCGAGCAACAGATCAGAAAGATAATTTGCTTTACAGAAATG 480
QY 481 GGGAGCGATCTCTGAAATCGCCATTCAGTCCGACTTGGATGCTTCCACTCGGATTTTC 540
DB 481 GGGAGCGATCTCTGAAATCGCCATTCAGTCCGACTTGGATGCTTCCACTCGGATTTTC 540
QY 541 AAACATGCGCCAAAGAAAGTCTTGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
DB 541 AAACATGCGCCAAAGAAAGTCTTGAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGG 600
QY 601 AGCCGGGTGTGCCAGCTGATCAACACTTGCAAGCCGTCGACCCAGTCTTCTTGGCCCA 660
DB 601 AGCCGGGTGTGCCAGCTGATCAACACTTGCAAGCCGTCGACCCAGTCTTCTTGGCCCA 660
QY 661 CCCCGCAGCTGTTGACTCAACAGTCCCTCTGAGCAAGGACACCGCGCTCCCTCGGCCG 720
DB 661 CCCCGCAGCTGTTGACTCAACAGTCCCTCTGAGCAAGGACACCGCGCTCCCTCGGCCG 720
QY 721 CCGGTCGCGCGCGCGCTTGCCTGAGAGCGCGCGGAGAGCTGGAGCCCTCGCT 780
DB 721 CCGGTCGCGCGCGCGCTTGCCTGAGAGCGCGCGGAGAGCTGGAGCCCTCGCT 780
QY 781 ACTGCGTGCCTGATCTCAGCGGACTCAGCCAGCCGCTGAGCTCGCGCGGAGCAAGCA 840
DB 781 ACTGCGTGCCTGATCTCAGCGGACTCAGCCAGCCGCTGAGCTCGCGCGGAGCAAGCA 840
QY 841 CGACACCGACAGCGCTACGCGGGAAGCGAGCGCGCGCGCGGAGCAAGCAAGCA 900
DB 841 CGACACCGACAGCGCTACGCGGGAAGCGAGCGCGCGCGCGGAGCAAGCAAGCA 900
QY 901 AAGCGCGGGGAGCGCGCTCACCATCAAGCAGGAGCTCCCGGGAGGACTCGCGCG 960
DB 901 AAGCGCGGGGAGCGCGCTCACCATCAAGCAGGAGCTCCCGGGAGGACTCGCGCG 960
QY 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGCGGCGGAGCGCGCGCGCGCGCGCGG 1020
DB 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGCGGCGGAGCGCGCGCGCGCGCGCGG 1020
QY 1021 CGCGCGCGGCGGAGCGCGCTTCTGGGCGCGGAGCCCTCGCGCGCGCGCGCGCTGC 1080
DB 1021 CGCGCGCGGCGGAGCGCGCTTCTGGGCGCGGAGCCCTCGCGCGCGCGCGCGCTGC 1080
QY 1081 TGAGACCCGACGCGCGCTGCTCAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 TGAGACCCGACGCGCGCTGCTCAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 CTTTCCCGAGCG 1200
DB 1141 CTTTCCCGAGCG 1200
QY 1201 CTTTCTGAGCTGCGCGCTACGTGACGCGCTTCTGCAAGAGAGCGCGCTGGAGAAATC 1260
DB 1201 CTTTCTGAGCTGCGCGCTACGTGACGCGCTTCTGCAAGAGAGCGCGCTGGAGAAATC 1260
QY 1261 TGTACCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
DB 1261 TGTACCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
QY 1321 CGGAGCG 1380
DB 1321 CGGAGCG 1380
QY 1381 CCTCGGTTGTCGCGCGCTTCCGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1381 CCTCGGTTGTCGCGCGCTTCCGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 ACAGGTTGGCGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
DB 1441 ACAGGTTGGCGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500

Qy	1501	CC	TTCCGGGCGGCGCGGAGCCGGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC	1560
Db	1501	CC	TTCCGGGCGGCGGCGGAGCCGGGGAACCGGAGAGCTCTGCTCAGGAAGATCCCTCGC	1560
Qy	1561	AG	CCAGGAAAGGAAGCTCCCTGGAATCCCTTGCCTCCCGAAGGACGGAGGTTCAAAGCAGAGT	1620
Db	1561	AG	CCAGGAAAGGAAGCTCCCTGGAATCCCTTGCCTCCCGAAGGACGGAGGTTCAAAGCAGAGT	1620
Qy	1621	GAG	AGGTTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAGTTAAATATACCCTTAAG	1680
Db	1621	GAG	AGGTTAAATACCCCTTAAGGAGGTTCAAGCAGAGTGAGAGTTAAATATACCCTTAAG	1680
Qy	1681	GT	CTTTAAGGAGGAAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG	1740
Db	1681	GT	CTTTAAGGAGGAAAGTGTAATAGATGCACGACAGGCATAAACAAGAAACAACAACAG	1740
Qy	1741	GT	GTATGTGTACATTCGGAGTTCCCTGTTTGTCTATCCGCGACACCCACCCCTCGACA	1800
Db	1741	GT	GTATGTGTACATTCGGAGTTCCCTGTTTGTCTATCCGCGACACCCACCCCTCGACA	1800
Qy	1801	CAC	TAAACATCCCTTTCTCCCCACCACAGCTGTAAAGATCCCTATGCGAAGACACTGGC	1860
Db	1801	CAC	TAAACATCCCTTTCTCCCCACCACAGCTGTAAAGATCCCTATGCGAAGACACTGGC	1860
Qy	1861	TC	TTTTTTTAAATCCCCCAATAAATTTTGCCTTTTATAGGCATGTTCCATATCTCT	1920
Db	1861	TC	TTTTTTTAAATCCCCCAATAAATTTTGCCTTTTATAGGCATGTTCCATATCTCT	1920
Qy	1921	TAA	ATTTGGAACCTTAATTCGAGAGGAAGTAAGAAGGTCGTGTTCTGTGGCTGAGCTAGGT	1980
Db	1921	TAA	ATTTGGAACCTTAATTCGAGAGGAAGTAAGAAGGTCGTGTTCTGTGGCTGAGCTAGGT	1980
Qy	1981	GA	CCCCGGGTAGGGAAAGATGTTAAACCTTTTGACGCTTTTGGAGTTGACATGGAAC	2040
Db	1981	GA	CCCCGGGTAGGGAAAGATGTTAAACCTTTTGACGCTTTTGGAGTTGACATGGAAC	2040
Qy	2041	AG	CAGGTAGTGTATCTAGAGCTAGTTCTCAAGCTGCCCTGCCCTGTTTATGGAGGCGT	2100
Db	2041	AG	CAGGTAGTGTATCTAGAGCTAGTTCTCAAGCTGCCCTGCCCTGTTTATGGAGGCGT	2100
Qy	2101	TC	CAACACAGATGAGGCTCTTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGT	2160
Db	2101	TC	CAACACAGATGAGGCTCTTTTAGAATTGAATTTACTCTTCAGTATTTTCTAATGT	2160
Qy	2161	TC	AGCTTTCAAGGCATATATTTTCAAGAGTAGAGATGCAGTTTCTCAGCTGCA	2220
Db	2161	TC	AGCTTTCAAGGCATATATTTTCAAGAGTAGAGATGCAGTTTCTCAGCTGCA	2220
Qy	2221	AC	CTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTGCACCTGTTTAAAGAAACAC	2280
Db	2221	AC	CTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTGCACCTGTTTAAAGAAACAC	2280
Qy	2281	GG	AGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAGGCCAGAAATCA	2340
Db	2281	GG	AGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGGATGAGAGGCCAGAAATCA	2340
Qy	2341	TT	CTAGTACATTTGCTTAAACACTTTATGAGAAATTGACCATGAAATTATGGACTCATCT	2400
Db	2341	TT	CTAGTACATTTGCTTAAACACTTTATGAGAAATTGACCATGAAATTATGGACTCATCT	2400
Qy	2401	TA	ATTTCTTCAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
Db	2401	TA	ATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA	2460
Qy	2461	TAG	ATAGTATCTATACATACATCATCAAGTGCATCTATTTCCCATCTCTCATTTAATCCAT	2520
Db	2461	TAG	ATAGTATCTATACATACATCATCTCAAGTGCATCTATTTCCCATCTCTCATTTAATCCAT	2520
Qy	2521	CA	TGTTCCCTAAATTTTGTAACTTTACTGTAAAAAAGGCGACTGAACCTCAAAACAAA	2580
Db	2521	CA	TGTTCCCTAAATTTTGTAACTTTACTGTAAAAAAGGCGACTGAACCTCAAAACAAA	2580

Qy	2581	ACAAAAACACACACACACAAACAAACACAGTCCAAACCTGATATATCCCTATATCTCGTTAAAA	2641
Db	2581	ACAAAAACACACACACAAACAAACAAACACAGTCCAAACCTGATATATCCCTATATCTCGTTAAAA	2640
Qy	2641	TTCAAAAGTGAACGAAAGCACTTTAACTGGCCAGCTTTTGAATGCAAAATGCTGTGAAGATAT	2700
Db	2641	TTCAAAAGTGAACGAAAGCACTTTAACTGGCCAGCTTTTGAATGCAAAATGCTGTGAAGATAT	2700
Qy	2701	AGAAATGAAGTCCTGTGAGGCGCTTCCCTATCTCCAAAGTCTATGTATTTTCTGGAGACCAAAAC	2760
Db	2701	AGAAATGAAGTCCTGTGAGGCGCTTCCCTATCTCCAAAGTCTATGTATTTTCTGGAGACCAAAAC	2760
Qy	2761	CAGATACCAAGATAAATCAACAAAGAAAGCTTTTAAATAAGGCTTTAAACCAAGACCTTGCTCT	2820
Db	2761	CAGATACCAAGATAAATCAACAAAGAAAGCTTTTAAATAAGGCTTTAAACCAAGACCTTGCTCT	2820
Qy	2821	AGATATTTTTAGTTGTTGTGCAAGGTAGCACTGTGAGAAATCTCACCTTGGATGTTATGTA	2880
Db	2821	AGATATTTTTAGTTGTTGTGCAAGGTAGCACTGTGAGAAATCTCACCTTGGATGTTATGTA	2880
Qy	2881	AGGGGTGAGACAAACAGTCTGACTATGAGTAGGAGAAATATCTGGGCTCTTTTCGTCAGT	2940
Db	2881	AGGGGTGAGACAAACAGTCTGACTATGAGTAGGAGAAATATCTGGGCTCTTTTCGTCAGT	2940
Qy	2941	TTGGTGCAATTTGCTGCTGCTGTTGCTACTGCTTTGCCCTCAACCGCTGTGTTTAAACAACGT	3000
Db	2941	TTGGTGCAATTTGCTGCTGCTGTTGCTACTGCTTTGCCCTCAACCGCTGTGTTTAAACAACGT	3000
Qy	3001	TAAACTCTTAGCCTCAACAGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATG	3060
Db	3001	TAAACTCTTAGCCTCAACAGTGGCTCTTATGTACATAGTTGTTAATACATCCAAATTAATG	3060
Qy	3061	ATGCTCAGCATGCTATTTTGTAGGAGAGAAATATGTCTAAATGATATTTTGAGTTAAAA	3120
Db	3061	ATGCTCAGCATGCTATTTTGTAGGAGAGAAATATGTCTAAATGATATTTTGAGTTAAAA	3120
Qy	3121	TATCTTTTGGGAGGATTTGCTGAAAAGTTGCATCTTTTGTACAAATGCTTATCTGGTA	3180
Db	3121	TATCTTTTGGGAGGATTTGCTGAAAAGTTGCATCTTTTGTACAAATGCTTATCTGGTA	3180
Qy	3181	CAAGCTTATGCTGCTCTTAAATTTTAAAAAAATTAATACTGCTGTGAGAGAACGAGC	3240
Db	3181	CAAGCTTATGCTGCTCTTAAATTTTAAAAAAATTAATACTGCTGTGAGAGAACGAGC	3240
Qy	3241	TGCTTTAGAAAAAGTTTAGTAGTAGGACGATAAATAGCAATTTACCTTTATATCTAGTATT	3300
Db	3241	TGCTTTAGAAAAAGTTTAGTAGTAGGACGATAAATAGCAATTTACCTTTATATCTAGTATT	3300
Qy	3301	TTCAGCACTCCATAAATCTTATTTACCTTAAATATTTGCCACACTATTTTGTGATTTAAAAAT	3360
Db	3301	TTCAGCACTCCATAAATCTTATTTACCTTAAATATTTGCCACACTATTTTGTGATTTAAAAAT	3360
Qy	3361	TCTTACTAAGGAATAAAACCTTTAATATACGATATGATATGTCTAATAATTTAAAAAGA	3420
Db	3361	TCTTACTAAGGAATAAAACCTTTAATATACGATATGATATGTCTAATAATTTAAAAAGA	3420
Qy	3421	CATAATGGATGCTCAATTAGTTTAAAGATATCTAATACTATAGGATACAAATCACTACA	3480
Db	3421	CATAATGGATGCTCAATTAGTTTAAAGATATCTAATACTATAGGATACAAATCACTACA	3480
Qy	3481	GTTCTCAGATTTACACTTTTTTTTGTCAATGCTGTGATGTCTACATTTTCCAAATCTCTTT	3540
Db	3481	GTTCTCAGATTTACACTTTTTTTTGTCAATGCTGTGATGTCTACATTTTCCAAATCTCTTT	3540
Qy	3541	GCAAGCCTCCAGGCTCTGGCTTTGCTCTACTGCTGCTTCCCAATGTATCTTAATGAAAAG	3600
Db	3541	GCAAGCCTCCAGGCTCTGGCTTTGCTCTACTGCTGCTTCCCAATGTATCTTAATGAAAAG	3600
Qy	3601	TGCAAAAGAAAAACCTACCAATTTAAAAAAAAAAAAAAAAAA	3641
Db	3601	TGCAAAAGAAAAACCTACCAATTTAAAAAAAAAAAAAAAAAA	3641

Db 1441 ACAGAGTGGCGCCCTTGGGGCGCGCACCCCAGCACCCGCGACCGCGCGCACCCACCTGC 1500
Qy 1501 CTTTCGCGCGCCCGCGAGCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTGCG 1560
Db 1501 CTTTCGCGCGCCCGCGAGCGGGGAAACCGGAGAGCTCTGCTCAGGAAGATCCCTGCG 1560
Qy 1561 AGCCAGAAAGAGAGCTCCCTGAAATCCTTCGCTCCGGAAGGAGGAGGTTCAAGCAGAGT 1620
Db 1561 AGCCAGAAAGAGAGCTCCCTGAAATCCTTCGCTCCGGAAGGAGGAGGTTCAAGCAGAGT 1620
Qy 1621 GAGAGGTAAATATACCCCTTAAGGAGGTTCAAGCAGAGTGAAGGTTTAAATACCCCTTAAG 1680
Db 1621 GAGAGGTAAATATACCCCTTAAGGAGGTTCAAGCAGAGTGAAGGTTTAAATACCCCTTAAG 1680
Qy 1681 GTCTTTAAGGAGGAGAGTGTAAATAGATGACGACAGGCAATAACAAAGAACAAACAAACAG 1740
Db 1681 GTCTTTAAGGAGGAGAGTGTAAATAGATGACGACAGGCAATAACAAAGAACAAACAAACAG 1740
Qy 1741 GTGTTATGTACATTCGAGGTTCCGTTTGTCTATCCCGCACCACCCACCCCTCCACA 1800
Db 1741 GTGTTATGTACATTCGAGGTTCCGTTTGTCTATCCCGCACCACCCACCCCTCCACA 1800
Qy 1801 CACTAACATCCCTTTCTCCCGCCACCCAGCTGTAAAGATCCTATCGGAAAGACACTGCG 1860
Db 1801 CACTAACATCCCTTTCTCCCGCCACCCAGCTGTAAAGATCCTATCGGAAAGACACTGCG 1860
Qy 1861 TCTTTTTTTAATCCCGCAATTAATTTGGCCCTTTTAGGCCATGTCCTATATCTCT 1920
Db 1861 TCTTTTTTTAATCCCGCAATTAATTTGGCCCTTTTAGGCCATGTCCTATATCTCT 1920
Qy 1921 TAAAAATGGAACTAATTCGAGAGGAGTAAGAGGGTCTGTTCTGCTGAGCTAGGT 1980
Db 1921 TAAAAATGGAACTAATTCGAGAGGAGTAAGAGGGTCTGTTCTGCTGAGCTAGGT 1980
Qy 1981 GAAACCCCGGGTAGGGGAAAGATTAACACCTTTGAGCTCTTTGAGGTTGACATGGAAC 2040
Db 1981 GAAACCCCGGGTAGGGGAAAGATTAACACCTTTGAGCTCTTTGAGGTTGACATGGAAC 2040
Qy 2041 ASCAGTAGTCTGTATGTAGAGCTAGTCTCAAGAGCTCCCTGCTGTTTAGGAGGCGT 2100
Db 2041 ASCAGTAGTCTGTATGTAGAGCTAGTCTCAAGAGCTCCCTGCTGTTTAGGAGGCGT 2100
Qy 2101 TCCACAAACAGATTGAGGCTCTTTTAGAATTAATTTACTCTTCAGTATTTCTAATGT 2160
Db 2101 TCCACAAACAGATTGAGGCTCTTTTAGAATTAATTTACTCTTCAGTATTTCTAATGT 2160
Qy 2161 TCAGCTTTCTAAGGCATATATTTTCAAGAGTGAGGATGAGTCTCAAGTTGCA 2220
Db 2161 TCAGCTTTCTAAGGCATATATTTTCAAGAGTGAGGATGAGTCTCAAGTTGCA 2220
Qy 2221 ACCTATTCGAGTGGTTTAAATGGTATCTCTTAGTAATCTGCACTCCGTTAAAGAAACAC 2280
Db 2221 ACCTATTCGAGTGGTTTAAATGGTATCTCTTAGTAATCTGCACTCCGTTAAAGAAACAC 2280
Qy 2281 GGAGCTGGCCATCTGAGAACTAAGTCAGGAGGAGATGGATGAGAGGCCAGAAATCA 2340
Db 2281 GGAGCTGGCCATCTGAGAACTAAGTCAGGAGGAGATGGATGAGAGGCCAGAAATCA 2340
Qy 2341 TTCCTAGTACATTTGCTAACACTTTATTGAGAAATTTGACCATGAAATTAATGCACTCT 2400
Db 2341 TTCCTAGTACATTTGCTAACACTTTATTGAGAAATTTGACCATGAAATTAATGCACTCT 2400
Qy 2401 TAAATTCCTTAAGTCCATATATAGATATCTATCTGACAGATTTCTATTATCCA 2460
Db 2401 TAAATTCCTTAAGTCCATATATAGATATCTATCTGACAGATTTCTATTATCCA 2460
Qy 2461 TAGATAGTATCTATACATACATCTCAAGTGCATCTATTTCCCACTCTCATTTCCAT 2520
Db 2461 TAGATAGTATCTATACATACATCTCAAGTGCATCTATTTCCCACTCTCATTTCCAT 2520
Qy 2521 CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGAGTGCATCTGAACTTCAAAACAAA 2580
Db 2521 CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGAGTGCATCTGAACTTCAAAACAAA 2580

Qy 2581 AAAAAAACAAACAAACAAACAAACAAAGTCCAACTGATATATCTATATCTCTTAAAA 2640
Db 2581 AAAAAAACAAACAAACAAACAAACAAAGTCCAACTGATATATCTATATCTCTTAAAA 2640
Qy 2641 TTCAAAAGTGAAAGCAATTTAACTGGCCAGTCTTTGATGTCAAATGCTGTAAAGATAT 2700
Db 2641 TTCAAAAGTGAAAGCAATTTAACTGGCCAGTCTTTGATGTCAAATGCTGTAAAGATAT 2700
Qy 2701 AGAATGAAGTCTCTGTAGGCTCTCTATCTCCAAAGTCTATGTTCTTCGAGACCAAC 2760
Db 2701 AGAATGAAGTCTCTGTAGGCTCTCTATCTCCAAAGTCTATGTTCTTCGAGACCAAC 2760
Qy 2761 CAGATACAGATATACAAAGAAAGCTTTTTTAATAGGCTTTAAACCAAGACCTTGTCT 2820
Db 2761 CAGATACAGATATACAAAGAAAGCTTTTTTAATAGGCTTTAAACCAAGACCTTGTCT 2820
Qy 2821 AGATATTTTGTAGTTTGTGCAAGGTAGCACTGTGAGAAATCTCAGTTGGATGTTATGTA 2880
Db 2821 AGATATTTTGTAGTTTGTGCAAGGTAGCACTGTGAGAAATCTCAGTTGGATGTTATGTA 2880
Qy 2881 AGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAATATCTGGTCTTTTCGTCAGT 2940
Db 2881 AGGGGTGAGACACAAACAGTCTGACTATGAGTGAGGAAATATCTGGTCTTTTCGTCAGT 2940
Qy 2941 TTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTGGCTCAAAAGCTGTGTTTAAACACGT 3000
Db 2941 TTGGTGCAATTTGCTGCTGCTGTTGCTACTGTTGGCTCAAAAGCTGTGTTTAAACACGT 3000
Qy 3001 TAAACTCTTAGCTTACAAAGGTGCTTATGTACATAGTTGTTAAATACATCCCAATTAATG 3060
Db 3001 TAAACTCTTAGCTTACAAAGGTGCTTATGTACATAGTTGTTAAATACATCCCAATTAATG 3060
Qy 3061 ATGTCTGACATGCTATTTTGTAGGAGAAATATGTGCTPAATGATATTTTGAGTTAAAA 3120
Db 3061 ATGTCTGACATGCTATTTTGTAGGAGAAATATGTGCTPAATGATATTTTGAGTTAAAA 3120
Qy 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTGCACTTTGTTTCAATGCTTATCTCTGTTA 3180
Db 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTGCACTTTGTTTCAATGCTTATCTCTGTTA 3180
Qy 3181 CAAGCTTATGCTCTTTTAAATTTTAAATAATTTAAATACTGCTGTGAGAAACACAGC 3240
Db 3181 CAAGCTTATGCTCTTTTAAATTTTAAATAATTTAAATACTGCTGTGAGAAACACAGC 3240
Qy 3241 TGGTTTGAAGAAAGTTAGTATGACGATAGAAATAGAAATAGCTTTATATCTTAGTATT 3300
Db 3241 TGGTTTGAAGAAAGTTAGTATGACGATAGAAATAGAAATAGCTTTATATCTTAGTATT 3300
Qy 3301 TTGAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGATTTAAAAAT 3360
Db 3301 TTGAGCACTCCATAAATTTCTATTACCTAAATATTGCCACACTATTTTGATTTAAAAAT 3360
Qy 3361 TCTTACTAAGAAATAAAACCTTTAATATAGATATGATTTGCTAATAATTTAAAAAAGA 3420
Db 3361 TCTTACTAAGAAATAAAACCTTTAATATAGATATGATTTGCTAATAATTTAAAAAAGA 3420
Qy 3421 CATAATGATGCTCAATTAGTTTTTAAGATATCTATACTATAGGATACAAATCACTACA 3480
Db 3421 CATAATGATGCTCAATTAGTTTTTAAGATATCTATACTATAGGATACAAATCACTACA 3480
Qy 3481 GTTCTCAGATTTACACCTTTTTTGTCAATGAGCTTTGATGTCACACATTTCCAATCTCT 3540
Db 3481 GTTCTCAGATTTACACCTTTTTTGTCAATGAGCTTTGATGTCACACATTTCCAATCTCT 3540
Qy 3541 GCAAGCTCCAGCTCTGGCTTTGCTACTGCTCGTTCCCAATGATCTTAATGAAAG 3600
Db 3541 GCAAGCTCCAGCTCTGGCTTTGCTACTGCTCGTTCCCAATGATCTTAATGAAAG 3600
Qy 3601 TGCAAAAGAAAAACCTTACCAATTTAAAAAAGAAAAAAGAAAAA 3641
Db 3601 TGCAAAAGAAAAACCTTACCAATTTAAAAAAGAAAAAAGAAAAA 3641

RESULT 4
ADN05013
ID ADN05013 standard; cDNA; 3641 BP.
XX
AC ADN05013;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #721.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN W02004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
XX WPI; 2004-305105/28.
DR P-PSDB; ADN05014.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
XX Claim 1; SEQ ID NO 1407; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 3641 BP; 994 A; 922 C; 836 G; 889 T; 0 U; 0 Other;
Query Match 100.0%; Score 3641; DB 12; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCACTGAAGAGGGAGAGCGAGAGAGACTGGAGACGCACAGATCCCCCAAGTCTC 60
DB 1 CTGCACTGAAGAGGGAGAGCGAGAGAGACTGGAGACGCACAGATCCCCCAAGTCTC 60
QY 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCAAAAATCGAAACAGAGAAACG 120
DB 61 CCAAGCCTACCGTCCACAGATATTGTACAGAGCCCAAAAATCGAAACAGAGAAACG 120
QY 121 AACAGCAGTTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGATTACTGGAAC 180
DB 121 AACAGCAGTTGAACATGGACGAAGGAATTCCTCATTTGCAAGAGAGACAGATTACTGGAAC 180
QY 181 ATAGAGATTTTATAGACTGGACTATTCCTCTTTTGTATATGTATATGTAAACCCAAAGAGCA 240
DB 181 ATAGAGATTTTATAGACTGGACTATTCCTCTTTTGTATATGTATATGTAAACCCAAAGAGCA 240
QY 241 TGAACCGAGACGACCAAGGATPACCTACAAATTTACCGCACAGATTATATAGAAAAAGAAA 300
DB 241 TGAACCGAGACGACCAAGGATPACCTACAAATTTACCGCACAGATTATATAGAAAAAGAAA 300
QY 301 GAAGAGACCGAATTAATGAATGATTCAGCTGAAAGATTACTGCTGAAATCTCTGA 360
DB 301 GAAGAGACCGAATTAATGAATGATTCAGCTGAAAGATTACTGCTGAAATCTCTGA 360

QY 361 AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACCTTTGAACACT 420
DB 361 AATTGACAACTCTGGGACATCTGGAGAAAAGCTGTAGTCTTGGAAATTAACCTTTGAACACT 420
QY 421 TAAAGCTTTTAAACGCTTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATG 480
DB 421 TAAAGCTTTTAAACGCTTTAAACCGAGCAACAGCATCAGAAAGATAATTGCTTTACAGAATG 480
QY 481 GGGAGCGATCTCTGAAATCGCCCAATTCAGTCCGACTTTGAGATGCTTCCTCGGATTTTC 540
DB 481 GGGAGCGATCTCTGAAATCGCCCAATTCAGTCCGACTTTGAGATGCTTCCTCGGATTTTC 540
QY 541 AAACATGCGCCAAAGAGTCTTGCATACCTCTCCCGGTTTGGAGCTCGACACCCAGGG 600
DB 541 AAACATGCGCCAAAGAGTCTTGCATACCTCTCCCGGTTTGGAGCTCGACACCCAGGG 600
QY 601 AGCCGCGGTGTCTCCAGCTGATCAACCACTTGACGCCGCTGCCACCCAGTTCTTTCGCCA 660
DB 601 AGCCGCGGTGTCTCCAGCTGATCAACCACTTGACGCCGCTGCCACCCAGTTCTTTCGCCA 660
QY 661 CCCGCGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCAACCGCGCTCCCTCGGCG 720
DB 661 CCCGCGAGCTGTTGACTCAACAGGTCCCTCTGAGCAAGGCAACCGCGCTCCCTCGGCG 720
QY 721 CCGGTCGCGCGCGCGCCCTCTGCTGAGCGCGCGGCGAGAGCTGGAGCCCTCGCT 780
DB 721 CCGGTCGCGCGCGCGCCCTCTGCTGAGCGCGCGGCGAGAGCTGGAGCCCTCGCT 780
QY 781 ACTGCTGCTCCCTCATCTCAGCGGACTCAGCCAGCCCGAGCTGCGCGCGGAGACGACA 840
DB 781 ACTGCTGCTCCCTCATCTCAGCGGACTCAGCCAGCCCGAGCTGCGCGCGGAGACGACA 840
QY 841 CGGACACCGAGCGGTACGCGCGGCGAGAGCCCGCGCGGAGCCCGCGGAGAGGCA 900
DB 841 CGGACACCGAGCGGTACGCGCGGCGAGAGCCCGCGCGGAGCCCGCGGAGAGGCA 900
QY 901 AAGCGCGGGCGAGCCGCTCACCATCAAGCAGAGCTCTCCGCGGAGAGACTCGCCCG 960
DB 901 AAGCGCGGGCGAGCCGCTCACCATCAAGCAGAGCTCTCCGCGGAGAGACTCGCCCG 960
QY 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGCGCGGCGAGCGCGCGCGCGCGCGCGG 1020
DB 961 CGCCCAAGAGGATGAAGCTGGATTCCCGGCGCGGCGAGCGCGCGCGCGCGCGCGG 1020
QY 1021 GCGCGCGCGCGCGAGCGCTCTGCGGCGCGGAGCCCTCTGCGCGCGCGCGCGCTGC 1080
DB 1021 GCGCGCGCGCGCGAGCGCTCTGCGGCGCGGAGCCCTCTGCGCGCGCGCGCGCTGC 1080
QY 1081 TGAGACCCGAGCGCGCTGCTCAGCTCGCTGTGTGCTGCTGCGGAGGAGGCGCG 1140
DB 1081 TGAGACCCGAGCGCGCTGCTCAGCTCGCTGTGTGCTGCTGCGGAGGAGGCGCG 1140
QY 1141 CTTTCCGCGAGCGCGCGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 CTTTCCGCGAGCGCGCGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 CTTTCTGAGCTGCGCGCTACGTGACGCTTCTGCAAGAGCGCGCTTGAGAAATATC 1260
DB 1201 CTTTCTGAGCTGCGCGCTACGTGACGCTTCTGCAAGAGCGCGCTTGAGAAATATC 1260
QY 1261 TGTACCCGCGCGCGCTGCGCGCGCTTCCGCTGTATACCCCGCATCCCCCGCGCG 1320
DB 1261 TGTACCCGCGCGCGCTGCGCGCGCTTCCGCTGTATACCCCGCATCCCCCGCGCG 1320
QY 1321 CGGCGCGCGCGAGCGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCTGCTGCTGCTG 1380
DB 1321 CGGCGCGCGCGAGCGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCTGCTGCTGCTG 1380
QY 1381 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 ACAGAGTGGCGCGCTTGGCGCGCGCGCACCCCGAGCACCGCGCGCGCGCGCGCTGC 1500

QY 901 AAGCGCGGGGCGAGCGCGCTCACCATCAAGCAGAGAGCTCCCGGGAGGAGCTGCGCGG 960
Db 1017 AAGCGCGGGGCGAGCGCGCTCACCATCAAGCAGAGAGCTCCCGGGAGGAGCTGCGCGG 1076
QY 961 CGCCCAAGAGCATGAAGCTGGATTCCGGGGCGGCGAGCGCGCGCGCGCGCGCGG 1020
Db 1077 CGCCCAAGAGCATGAAGCTGGATTCCGGGGCGGCGAGCGCGCGCGCGCGCGG 1136
QY 1021 GCGCGGGGCGGCGAGCGCGCTTCTGGGGCCGACCTTCCGCGCGCGCGCGCGCTGC 1080
Db 1137 GCGCGGGGCGGCGAGCGCGCTTCTGGGGCCGACCTTCCGCGCGCGCGCGCGCTGC 1196
QY 1081 TGAGACCCGAGCGCGCTGCTCAGCTCGCTGGTGGCTTCCGCGAGCGAGCGCGCG 1140
Db 1197 TGAGACCCGAGCGCGCTGCTCAGCTCGCTGGTGGCTTCCGCGAGCGAGCGCGCG 1256
QY 1141 CTTTCCCGCAGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTCTCGC 1200
Db 1257 CTTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCGCTTCTGCTTCTCTCGC 1316
QY 1201 CTTTCTGAGCTGCGGCTAGTGCAGCCCTTCTTGACAAGAGCGGCTTGAGAGATTC 1260
Db 1317 CTTTCTGAGCTGCGGCTAGTGCAGCCCTTCTTGACAAGAGCGGCTTGAGAGATTC 1376
QY 1261 TGTACC CGCGCGGCTGCGCGCGCTTCCCGCTGCTATACCGCGCATCCCGCGCGG 1320
Db 1377 TGTACC CGCGCGGCTGCGCGCGCTTCCCGCTGCTATACCGCGCATCCCGCGCGG 1436
QY 1321 CGCAGCGCGGCGAGCGCGCGCGCTGCCGCGCGCGCGCGCGCTTCCCTGCTGCT 1380
Db 1437 CGCAGCGCGGCGAGCGCGCGCGCTGCCGCGCGCGCGCGCGCTTCCCTGCTGCT 1496
QY 1381 CTTCCGCTGCTGCGCGCTCCGAGAGCGGCGCGCGCGCGCGCGCTTCCCTGCGCG 1440
Db 1497 CTTCCGCTGCTGCGCGCTCCGAGAGCGGCGCGCGCGCGCGCGCGCTTCCCTGCGCG 1556
QY 1441 ACGAGTGGCGCGCTTGGGGCGCGCACCCCGCAGCACCGCGCGCGCGCGCGCGCTG 1500
Db 1557 ACGAGTGGCGCGCTTGGGGCGCGCACCCCGCAGCACCGCGCGCGCGCGCGCGCTG 1616
QY 1501 CTTTCCGCGCGCGCGCGAGCGGGAACCGGAGCTCTGCTCAGAGAGATCCCTGCG 1560
Db 1617 CTTTCCGCGCGCGCGCGAGCGGGAACCGGAGCTCTGCTCAGAGAGATCCCTGCG 1676
QY 1561 AGCCAGGAAGAGCTCCTGAATCCTTGCTTCCGAGGAGCGAGGTTCAAGCAGAGT 1620
Db 1677 AGCCAGGAAGAGCTCCTGAATCCTTGCTTCCGAGGAGCGAGGTTCAAGCAGAGT 1736
QY 1621 GAGAAGTTAAATACCTTTAAGAGGTTCAAGCAGAGTGAGAGTTAAATACCTTTAAG 1680
Db 1737 GAGAAGTTAAATACCTTTAAGAGGTTCAAGCAGAGTGAGAGTTAAATACCTTTAAG 1796
QY 1681 GTCTTTAAGGAGGAGTGAATAGATGACACAGGAGTAAACAGAGCAACAAACAG 1740
Db 1797 GTCTTTAAGGAGGAGTGAATAGATGACACAGGAGTAAACAGAGCAACAAACAG 1856
QY 1741 GTGTTATGTACATTCGAGTTCCTGTTTCTCATCCCGCACCAACCCCTCCACA 1800
Db 1857 GTGTTATGTACATTCGAGTTCCTGTTTCTCATCCCGCACCAACCCCTCCACA 1916
QY 1801 CACTAATACCTTCTTCCCGCAGCAGCTGTAAGATCCTATGCGAAGACACTGCG 1860
Db 1917 CACTAATACCTTCTTCCCGCAGCAGCTGTAAGATCCTATGCGAAGACACTGCG 1976
QY 1861 TCTTTTTTTTAAATCCCGCAATAAATTTTGGCCCTTTTAGGCCATGTTCCATATCTCT 1920
Db 1977 TCTTTTTTTTAAATCCCGCAATAAATTTTGGCCCTTTTAGGCCATGTTCCATATCTCT 2036
QY 1921 TAAATTTGAACTTAATCCGAGGAGTAAAGAGGCTGTTCTGTGCGCTGAGCTAGT 1980
Db 2037 TAAATTTGAACTTAATCCGAGGAGTAAAGAGGCTGTTCTGTGCGCTGAGCTAGT 2096
QY 1981 GAACCCCGGGTAGGGGAAGATGTTACACCTTTGACGCTTTTGGAGTTGACATGGAAC 2040

Db 2097 GAACCCCGGGTAGGGGAAGATGTTAAACCTTTGACGCTTTTGGAGTTGACATGGAAC 2156
QY 2041 AGCAGGTAGTTGTTATATGATGAGCTAGTCTCAAGCTGCGCTGCTTTTAGAGCGGT 2100
Db 2157 AGCAGGTAGTTGTTATGATGAGCTAGTCTCAAGCTGCGCTGCTTTTAGAGCGGT 2216
QY 2101 TCCAAACACAGATTGAGGCTCTTTTAAAGTTGAATTTACTCTTCAGTATTTTCTAATGT 2160
Db 2217 TCCAAACACAGATTGAGGCTC-TTTTAAAGTTGAATTTACTCTTCAGTATTTTCTAATGT 2275
QY 2161 TCAGCTTTCTTAAAGGCATATATTTTCAAGAAGTGCAGATGCAGTTTCTCAGTTGCA 2220
Db 2276 TCAGCTTTCTTAAAGGCATATATTTTCAAGAAGTGCAGATGCAGTTTCTCAGTTGCA 2335
QY 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAACAC 2280
Db 2336 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAACAC 2395
QY 2281 GGAGCTGGCGCATCTGTCAGAACTAAGTCAGGAAGAGATGGATGAGAAGCCAGAAATCA 2340
Db 2396 GGAGCTGGCGCATCTGTCAGAACTAAGTCAGGAAGAGATGGATGAGAAGCCAGAAATCA 2455
QY 2341 TTCTCTAGTACATTTGCTAAACACTTTTATGAGAAATTCACCATGAATTAATGGACTATCT 2400
Db 2456 TTCTCTAGTACATTTGCTAAACACTTTTATGAGAAATTCACCATGAATTAATGGACTATCT 2515
QY 2401 TAAATTTCTTAAGTCCCATATATAGATAGATATCTCTGTACAGATTTCTTATTTATCCA 2460
Db 2516 TAAATTTCTTAAGTCCCATATATAGATAGATATCTCTGTACAGATTTCTTATTTATCCA 2575
QY 2461 TAGATAGTATCTATACATACACATCTCAAGTGCATCTATCCACCTCTCATTAATCCAT 2520
Db 2576 TAGATAGTATCTATACATACACATCTCAAGTGCATCTATCCACCTCTCATTAATCCAT 2635
QY 2521 CATGTTCTTAAATTTTGTAACTTACTGTAAATAAGTGCACCTGAACCTCAAAACAAA 2580
Db 2636 CATGTTCTTAAATTTTGTAACTTACTGTAAATAAGTGCACCTGAACCTCAAAACAAA 2695
QY 2581 ACAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2640
Db 2696 ACAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2755
QY 2641 TTTCAAAAGTGAAGAAAGCATTTAACTGCGCGAGTTTGTGTTGCAAAATGCTGTAAGATAT 2700
Db 2756 TTTCAAAAGTGAAGAAAGCATTTAACTGCGCGAGTTTGTGTTGCAAAATGCTGTAAGATAT 2815
QY 2701 AGAATGAAGTCTGTGAGGCGCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2760
Db 2816 AGAATGAAGTCTGTGAGGCGCTTCTATCTCCAAGTCTATGTATTTTCTGGAGACCAAC 2875
QY 2761 CAGTACAGATTAATCAAAAGAAAGCTTTTAAATAAGCTTTAAACCAAGACCTTGTCT 2820
Db 2876 CAGTACAGATTAATCAAAAGAAAGCTTTTAAATAAGCTTTAAACCAAGACCTTGTCT 2935
QY 2821 AGATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2880
Db 2936 AGATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2995
QY 2881 AGGGGTGAGACACAACAGTCTGATGAGTGAAGAAATATCTGGGCTCTTTTCTGTCAGT 2940
Db 2996 AGGGGTGAGACACAACAGTCTGATGAGTGAAGAAATATCTGGGCTCTTTTCTGTCAGT 3055
QY 2941 TTGTGTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Db 3056 TTGTGTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3115
QY 3001 TAAACTCTTAGCCTACAGGTGCTCTTATGTACATAGTTGTTTAAATACATCCATTAATG 3060
Db 3116 TAAACTCTTAGCCTACAGGTGCTCTTATGTACATAGTTGTTTAAATACATCCATTAATG 3175
QY 3061 ATGCTGACATGCTATTTTGTAGGGAGAAATATGCTGCTAATGATATTTTGTAGTAAAA 3120

Db	3176	ATGCTCAGATGCTATTTTTTGTAGGAGAAATAATGCTGCTAATGATATATTTTGAGTTAAA	3235
QY	3121	TATCTTTTGGGGAGGATTTGCTGAAAGATTGCATTTTGTTCACATGCTTAATGCTTGTA	3180
Db	3236	TATCTTTTGGGGAGGATTTGCTGAAAGATTGCATTTTGTTCACATGCTTAATGCTTGTA	3295
QY	3181	CAAGCTTATGCTGCTTAAATATTTTAAATAATTAATATCTGCTGAGAAACGAC	3240
Db	3296	CAAGCTTATGCTGCTTAAATATTTTAAAAAATAATATCTGCTGAGAAACGAC	3354
QY	3241	TGGTTTGAAGAAAGTTTAGTATGTGACGATAAATAGAAATTACCTTTATTTCTAGTATT	3300
Db	3355	TGGTTTGAAGAAAGTTTAGTATGTGACGATAAATAGAAATTACCTTTATTTCTAGTATT	3414
QY	3301	TTCCAGACTCCATAAATTCATTACCTTAATATTTGCCACACTATTTTGTGATTTAAAAT	3360
Db	3415	TTCCAGACTCCATAAATTCATTACCTTAAATATTTGCCACACTATTTTGTGATTTAAAAT	3474
QY	3361	TCCTTACTAAGGAATAAAAACTTTTAATATACGATA	3394
Db	3475	TCCTTACTAAGGAATAAAAACTTTTAATATACGATA	3508

RESULT 6

RESOLUTION 9
ABT05457
ID ABT05457 standard; DNA: 2304 BP.

AC ABT05457:

11-OCT-2002 (first entry)

XX
DE
DNA OF NOVX 5 SEQ TD NO 9

xx Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antiphaemic;
 KW neuroprotective; antiparkinsonian; anticoagulant; cerebroprotective;
 KW tranquilliser; neuroleptic; antidiabetic; antitumor; antinflammatory;
 KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOX; diabetes;
 KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
 KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
 KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
 KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
 KW vascular transport; cystic fibrosis; gastrointestinal disorder;
 KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
 KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
 KW gene therapy; gene; ds.

Unidentified.

XX
PN
W0300246409-A2

XX
13-JTN-200206-DEC-2001. 2001WC-IIS046586
XX
PF

06-DEC-2000. 2000IS-0251666P

PR 12-DEC-2000; 2000US-0255029P.
 08 JAN 2001 2001US-0260323P
 09 JAN 2001 2001US-0260323P

PR 24-JAN-2001; 2001US-0263800P.

PR 24-APR-2001; 2001US-0286183P.

PR 12-SEP-2001; 2001US-0318712P.

PA (CURA-) CURAGEN CORP.

PI Guo X, Li L, Patturajan M,

PI Edinger S, Peyman JA, Stone

XX

DR P-PSDB; ABJ04644.

XX PT Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic neurodegenerative, immune and hematopoietic disorders.

XX
PS
Claim 9. Page 43-44: 421pp: English:

The invention relates to an isolated polypeptide, designated NOVX, comprising a sequence fully defined in the specification. The isolated protein, its encoding polynucleotide or an antibody created from the protein is useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder, or for treating or preventing a NOVX-associated disorder in a subject, preferably human. The isolated protein, its encoding polynucleotide or an antibody created from the protein are also useful for treating or preventing metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder, immune disorders, haematopoietic disorders, and various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X, wasting disorders associated with chronic diseases, and cancer. The isolated protein, its encoding polynucleotide or an antibody created from the protein are useful for treating or preventing neurological disorders such as epilepsy, stroke, mental disorders including mood, anxiety, schizophrenic disorders, disorders of vesicular transport such as cystic fibrosis, diabetes mellitus, goiter, gastrointestinal disorders including ulcerative colitis, other conditions associated with abnormal vesicle trafficking including AIDS, allergic reactions, multiple sclerosis and rheumatoid arthritis. A cell comprising the vector of the invention is useful for producing non-human transgenic animals. The polynucleotide of the invention can be used to treat disorders by gene therapy. This polynucleotide sequence represents the DNA encoding one of the isolated NOVX proteins of the invention.

Sequence 2304 BP: 468 A: 784 C: 689 G: 363 T: 0 U: 0 Other: 0

Query Match 43.5%; Score 1583; DB 6; Length 2304;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1583; Conservative 0; Mismatches 0; Indels 0

1	QY	CTGCACCTGAAGAGGGAGAGCGGAGAGAGAGACTGGAGAGCGCACAGATCCCCCAGAGTCTC	60
722	Db	CTGCACCTGAAGAGGGAGAGCGGAGAGAGAGACTGGAGAGCGCACAGATCCCCCAGAGTCTC	781
61	QY	CCAAAGCCTTACCGTCCCGCCACAGATTATTGTACAGAGCCCAAAAATCGAAACAGAGAGAAACG	120
782	Db	CGAAGCCTTACCGTCCCGCCACAGATTATTGTACAGAGCCCAAAAATCGAAACAGAGAGAAACG	841
121	QY	AACGACGAGTTGAACATCGGACGAAGGAATTCCTCATTTGCAAGAGAGAGACAGTTACTCGGAAC	180
842	Db	AACGACGAGTTGAACATCGGACGAAGGAATTCCTCATTTGCAAGAGAGAGACAGTTACTCGGAAC	901
181	QY	ATAGAGATTTTATAGGACTCGGACTATTCCTCTTTGTATATGTGTAAACCCAAAAAGGAGCA	240
902	Db	ATAGAGATTTTATAGGACTCGGACTATTCCTCTTTGTATATGTGTAAACCCAAAAAGGAGCA	961
241	QY	TGAAACGAGAGCGACACCAAGGATACCTCACAAATTACCGCACAGATTAATAGAAAAGAAAA	300
962	Db	TGAAACGAGAGCGACACCAAGGATACCTCACAAATTACCGCACAGATTAATAGAAAAGAAAA	1021
301	QY	GAAGGAGCCGAAATTAATGAATGCATTGCTTCAGCTGAAAGATTTACTGCCTGAAACATCTGA	360
1022	Db	GAAGGAGCCGAAATTAATGAATGCATTGCTTCAGCTGAAAGATTTACTGCCTGAAACATCTGA	1081
361	QY	AATTGACAACCTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAAATTAACTTTGAAGACACT	420
1082	Db	AATTGACAACCTCTGGGACATCTGGAGAAAGCTGTAGTCTTTGGAAATTAACTTTGAAGACACT	1141
421	QY	TAAAGCTTTTAACCGCCCTTAAACGAGGAACGACATCAGAAAGATAATTGCTTTACAGAAATG	480
1142	Db	TAAAGCTTTTAAACCGCCCTTAAACGAGGAACGACATCAGAAAGATAATTGCTTTACAGAAATG	1201
481	QY	GGGAGCGATCTCTGAAATCGCCCAATTCAGTCGCGACTTGGATCGCTTCACACTCGGGATTTTC	540

Db 1202 GGGAGCGATCTCTGAATCGCCATTCAGTCCGATCGATGATGCTTCCATCTCGGATTC 1261
QY 541 AAACATCGCCCAAGAGTCTTCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGG 600
Db 1262 AAACATCGCCCAAGAGTCTTCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGG 1321
QY 601 AGCCGGGTGTCTCAGCTGATCAACACTTGACCCCGTGGCCACCCAGTCTTGGCCA 660
Db 1322 AGCCGGGTGTCTCAGCTGATCAACACTTGACCCCGTGGCCACCCAGTCTTGGCCA 1381
QY 661 CCCGAGCTGTGACTCAACAGTCCCTGAGCAAGGACCGGCTCCCTCGCCG 720
Db 1382 CCCGAGCTGTGACTCAACAGTCCCTGAGCAAGGACCGGCTCCCTCGCCG 1441
QY 721 CCGGTCGCGCGCCCTGCTGAGCGCGCGGCGAGAGCTGGAGCCCTCGCT 780
Db 1442 CCGGTCGCGCGCCCTGCTGAGCGCGCGGCGAGAGCTGGAGCCCTCGCT 1501
QY 781 ACTGCTGCCGTCATCCAGCGACTCAGCCAGCGCGAGCTGGCGCGGAGAGCA 840
Db 1502 ACTGCTGCCGTCATCCAGCGACTCAGCCAGCGCGAGCTGGCGCGGAGAGCA 1561
QY 841 CGACACCGACAGCGCTACGCGCGCAAGCCGCGCGAGCGCGCGGAGAGCA 900
Db 1562 CGACACCGACAGCGCTACGCGCGCAAGCCGCGCGAGCGCGCGGAGAGCA 1621
QY 901 AAGCGCGGCGGCGAGCGCGTCAACATCAGCAGAGCTCCCGGAGAGACTCGCCGG 960
Db 1622 AAGCGCGGCGGCGAGCGCGTCAACATCAGCAGAGCTCCCGGAGAGACTCGCCGG 1681
QY 961 CGCCCAAGAGATGAGCTGGAATCCCGCGGCGGCGAGCGCGCGCGCGCGGCG 1020
Db 1682 CGCCCAAGAGATGAGCTGGAATCCCGCGGCGGCGAGCGCGCGCGCGCGGCG 1741
QY 1021 GCGCGCGCGCGGCGAGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCGCTGC 1080
Db 1742 GCGCGCGCGCGGCGAGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCGCTGC 1801
QY 1081 TGAGACCGACGCGCGCTGCTAGCTCGCTGTGTGCGTTCGCGGAGCGAGCGCGCG 1140
Db 1802 TGAGACCGACGCGCGCTGCTAGCTCGCTGTGTGCGTTCGCGGAGCGAGCGCGCG 1861
QY 1141 CTTTCCCGCAGCGCGCGCGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCTGC 1200
Db 1862 CTTTCCCGCAGCGCGCGCGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCTGC 1921
QY 1201 CTTTCTGAGCTGCGCGCTACGTGAGCGCGCTTCTGAGCAAGAGCGCGCTTGGAGAGTATC 1260
Db 1922 CTTTCTGAGCTGCGCGCTACGTGAGCGCGCTTCTGAGCAAGAGCGCGCTTGGAGAGTATC 1981
QY 1261 TGTACCGCGCGCGCGCTGCGCGCTTCCCGCTGTATACCGCGGCTCCCGCGCGCG 1320
Db 1982 TGTACCGCGCGCGCGCTGCGCGCTTCCCGCTGTATACCGCGGCTCCCGCGCGCG 2041
QY 1321 CGGACCGCGCGCGCGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCTTCCCTGCTGT 1380
Db 2042 CGGACCGCGCGCGCGCGCGCTTCTGCGGCGCGCGCGCGCGCGCGCTTCCCTGCTGT 2101
QY 1381 CTTGCTGTGTGCGCGCTTCCGAGAGCGCGCGCGCGCGCGCGCGCTTCTGCGCG 1440
Db 2102 CTTGCTGTGTGCGCGCTTCCGAGAGCGCGCGCGCGCGCGCGCGCTTCTGCGCG 2161
QY 1441 ACGAGTGGCGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 1500
Db 2162 ACGAGTGGCGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 2221
QY 1501 CTTTCCCGCGCGCGCGCGCGCGAGCGCGGAGACCGGAGCTTCTCAGGAGATCCCTCG 1560
Db 2222 CTTTCCCGCGCGCGCGCGCGCGAGCGCGGAGACCGGAGCTTCTCAGGAGATCCCTCG 2281
QY 1561 ACCCAGGAGGAGCTCCCTCA 1583

Db 2282 AGCCAGGAAGGAGCTCCCTGA 2304
RESULT 7
ACD19338
ID ACD19338 standard; cDNA; 1471 BP.
XX
AC ACD19338;
XX
DT 25-AUG-2003 (first entry)
XX
DE cDNA encoding novel human protein #18.
XX
KW Human; NOV; gene therapy; endocrine related disease; diabetes;
metabolism-related disease; obesity; central nervous system disorder;
Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
schizophrenia; depression; autoimmune disorder; inflammatory disorder;
psoriasis; allergy; lupus erythematosus; asthma; cancer;
inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
lung disease; emphysema; obstructive pulmonary disease; haemophilia;
stroke; infection; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003023002-A2.
XX
PD 20-MAR-2003.
XX
PF 09-SEP-2002; 2002WO-US028539.
XX
PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 17-SEP-2001; 2001US-0322636P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 17-SEP-2001; 2001US-0322817P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325031P.
PR 26-SEP-2001; 2001US-0324990P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
Gerlach VL, Vernet CM, Ellerman K, Berghs C, Rothenberg WB, Guo X;
Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CB;
Rieger DK, Taupier RJ, Shenoy SG, Padigaru M, Alsobrook JP;
Lepley DM, Edinger SR, Burgess CE;
XX
WPI: 2003-313242/30.
XX
P-PSDB; ABO14645.
XX
PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
and polynucleotides, useful in gene therapy, e.g. for treating or
preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
stroke or infections.
XX
PS Claim 20; Page 125; 586pp; English.
XX
CC The invention describes a new isolated polypeptide (NOVX). The NOVX
polypeptide, nucleic acid and antibody are useful as therapeutics,
particularly in the manufacture of a medicament for treating a syndrome
associated with a human disease, which includes a pathology associated
with NOVX polypeptide. The DNA encoding the protein is useful in gene
therapy for treating the disease or condition. In particular, the NOVX
polypeptide or polynucleotide is useful for treating endocrine/

CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
 CC These are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications, and for monitoring the effects of drugs during clinical
 CC trials. This sequence encodes a novel human NOV protein
 XX
 SQ Sequence 1471 BP; 288 A; 515 C; 436 G; 232 T; 0 U; 0 Other;
 Query Match 40.4%; Score 1471; DB 10; Length 1471;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 129 TTGAAATGACGAGGAATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGA 248
 DB 1 TTGAACATGACGAGGAATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGA 60
 QY 189 TTTATAGGACTGACATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGA 248
 DB 61 TTTATAGGACTGACATTCCTCTTTGTATATGTGTAAACCCAAAGGAGCATGAAACGA 120
 QY 249 GACGACACCAAGGATACCTACAAATTCACCGACAGATTAATAGAAAGAAAGAGAGAC 308
 DB 121 GACGACACCAAGGATACCTACAAATTCACCGACAGATTAATAGAAAGAAAGAGAGAC 180
 QY 309 CGAATTAATGAATGCAATTCCTGAGCTGAAGATTTACTGCCTGAACATCTGAATTTGACA 368
 DB 181 CGAATTAATGAATGCAATTCCTGAGCTGAAGATTTACTGCCTGAACATCTGAATTTGACA 240
 QY 369 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCT 428
 DB 241 ACTCTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTGAACACACTTAAAGCT 300
 QY 429 TTAACCGCCTTAACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATCGGAGCGA 488
 DB 301 TTAACCGCCTTAACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATCGGAGCGA 360
 QY 489 TCTCTGAATCGCCATTCAGTCCGACTTGGATGCGTCCACTCGGATTTCAAAACATGC 548
 DB 361 TCTCTGAATCGCCATTCAGTCCGACTTGGATGCGTCCACTCGGATTTCAAAACATGC 420
 QY 549 GCCAAAGAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGAGCGCGG 608
 DB 421 GCCAAAGAGTCTTGCAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGAGCGCGG 480
 QY 609 TGTGTCAGCTGATCAACCACTTGACGCGCTGGCCACCCAGTCTTGCCACCCCGCGAG 668
 DB 481 TGTGTCAGCTGATCAACCACTTGACGCGCTGGCCACCCAGTCTTGCCACCCCGCGAG 540
 QY 669 CTGTTGACTCAACAGGTCCCTCTGAGCAAGGACCGGCGCTCCCTCGCGCCCGGGTCC 728
 DB 541 CTGTTGACTCAACAGGTCCCTCTGAGCAAGGACCGGCGCTCCCTCGCGCCCGGGTCC 600
 QY 729 GCGGCGCCCTCTGCTGAGCGCGGCGGACAGAGCTGGAGCCCTCGCTACTGCGTG 788
 DB 601 GCGGCGCCCTCTGCTGAGCGCGGCGGACAGAGCTGGAGCCCTCGCTACTGCGTG 660
 QY 789 CCGGTATCCAGCGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACGAGACAC 848
 DB 661 CCGGTATCCAGCGGACTCAGCCAGCGCGGAGCTCGCGCGGAGAACGACACGAGACAC 720
 QY 849 GACAGCGGCTACGCGCGGAGCGGAGCGCGCGGAGCGGAGAAAGGCAAGGCGG 908
 DB 721 GACAGCGGCTACGCGCGGAGCGGAGCGCGGAGCGGAGAAAGGCAAGGCGG 780
 QY 909 GGGGCGAGCGGCTACCATCAAGCAGGAGGAGCTCCCGGGGAGGAGCTCGCGCGCGCCCAAG 968

DB 781 GGGGCGAGCGCGCTACCATCAAGCAGGAGCTCCCGGAGGAGCTCGCGCGGCCCAAG 840
 QY 969 AGGATGAAGCTGATTCCTCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1028
 DB 841 AGGATGAAGCTGATTCCTCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
 QY 1029 GCGGCGGACGCGCGCTTCTGGGCGCGGACCTTGGCGCGCGCGCGCGCGCGCTGTGACACCC 1088
 DB 901 GCGGCGGACGCGCGCTTCTGGGCGCGGACCTTGGCGCGCGCGCGCGCGCGCTGTGACACCC 960
 QY 1089 GAGCGCGCGCTGTCTGACGCTGCTGGGCTTTCGGCGGAGGCGGAGCGCGCGCTTCCG 1148
 DB 961 GAGCGCGCGCTGTCTGACGCTGCTGGGCTTTCGGCGGAGGCGGAGCGCGCGCTTCCG 1020
 QY 1149 CAGCGCGCGCGCGCGCGCGCTTCTGCTGCTGCTTCTGCTTCTGCTGCTTCTGCA 1208
 DB 1021 CAGCGCGCGCGCGCGCGCGCTTCTGCTGCTGCTTCTGCTTCTGCTGCTTCTGCA 1080
 QY 1209 GCTGCGCGCTACGTCAGCGCGCTTCTTGGACAGAGCGGCTTGAGAAATATCTGTACCG 1268
 DB 1081 GCTGCGCGCTACGTCAGCGCGCTTCTTGGACAGAGCGGCTTGAGAAATATCTGTACCG 1140
 QY 1269 GCGCGCGCTGCGCGCGCTTCTGCTGCTATACCGCGGATCCCGCGCGCGCGCGCGCG 1328
 DB 1141 GCGCGCGCTGCGCGCGCTTCTGCTGCTATACCGCGGATCCCGCGCGCGCGCGCGCG 1200
 QY 1329 GCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCTTCCCTGCTGCTTCTGCTG 1388
 DB 1201 GCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCTTCCCTGCTGCTTCTGCTG 1260
 QY 1389 TGTGCGCGCGCTTCTGCGAGAGCGCGCGCGCGCGCGCGCTTCTGCTGCGACAGGCTG 1448
 DB 1261 TGTGCGCGCGCTTCTGCGAGAGCGCGCGCGCGCGCGCTTCTGCTGCGACAGGCTG 1320
 QY 1449 GCGCGCGCTTGGGCG 1508
 DB 1321 GCGCGCGCTTGGGCG 1380
 QY 1509 GGGCG 1568
 DB 1381 GGGCG 1440
 QY 1569 AAGGAAGCTCCCTGAATCCTTGGCTCCCGAA 1599
 DB 1441 AAGGAAGCTCCCTGAATCCTTGGCTCCCGAA 1471
 RESULT 8
 AAF74776
 ID AAF74776 standard; cDNA; 1511 BP.
 XX AAF74776;
 XX AAF74776;
 DT 17-MAY-2001 (first entry)
 XX
 DE Human DEC2b encoding cDNA sequence SEQ ID NO:11.
 XX
 KW DEC2a; DEC2b; bHLH type transcription factor; DEC2; DEC1;
 KW basic helix loop helix protein; cell differentiation; proliferation; ss.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 CDS 2..1456
 FT /*tag= a
 FT /product= "DEC2b"
 FT /note= "bHLH type transcription factor"
 XX
 PN WO200114551-A1.
 XX
 PD 01-MAR-2001.
 XX

PF 19-JUN-2000; 2000WO-JP003991.
XX
PR 19-AUG-1999; 99JP-00233286.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Fujimoto K, Shin M, Kato Y;
XX
XX WPI; 2001-202935/20.
DR P-PSDB; AAB70693.
XX
XX DEC2 is a basic helix loop helix protein of the DEC family for use in
PT development of drugs for treatment of disorders of cell differentiation
PT and proliferation.
XX
XX Claim 1; Page 63-68; 83pp; Japanese.
XX
XX The present invention describes a basic helix loop helix (bHLH) type
CC transcription factor designated DEC2. DEC2 can be used as a tool in the
CC development of drugs for the treatment and prevention of disorders
CC involving cell differentiation and proliferation. The present sequence
CC encodes the specifically claimed human DEC2b protein, as given in the
XX present invention
XX
SQ Sequence 1511 BP; 302 A; 521 C; 449 G; 239 T; 0 U; 0 Other;
Query Match 37.8%; Score 1378; DB 4; Length 1511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 261 GATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGACCGGAATTAATGAA 320
DB 134 GATACCTACAAATACCGACAGATTAATAGAAAAGAAAGAGACCGGAATTAATGAA 193
QY 321 TGCATTGCTCAGCTGAAAGATTTACTGCTGAACTCTGAACTTGAACACTTAAAGCTTTA 380
DB 194 TGCATTGCTCAGCTGAAAGATTTACTGCTGAACTCTGAACTTGAACACTTAAAGCTTTA 253
QY 381 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTA 440
DB 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACTTAAAGCTTTA 313
QY 441 ACCGAGCAACAGCATCAGAGATAATGCTTTACAGATGGGAGCGATCTCTGAATTCG 500
DB 314 ACCGAGCAACAGCATCAGAGATAATGCTTTACAGATGGGAGCGATCTCTGAATTCG 373
QY 501 CCCATTCACTCCGACTTGCATCGCTTCCACTCGGATTTCAACATGCGCAAGAGATC 560
DB 374 CCCATTCACTCCGACTTGCATCGCTTCCACTCGGATTTCAACATGCGCAAGAGATC 433
QY 561 TTGCAATACCTCTCCCGTTTGAAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 620
DB 434 TTGCAATACCTCTCCCGTTTGAAGCTGGACACCCAGGGAGCCCGGTGTGTCCAGCTG 493
QY 621 ATCAACCACTTGCAGCCGTGGCCACCCAGATTTCTGCCCACCCCGCATCTTGAATCAA 680
DB 494 ATCAACCACTTGCAGCCGTGGCCACCCAGATTTCTGCCCACCCCGCATCTTGAATCAA 553
QY 681 CAGGTCCTCTGAGCAAAAGGACACCGCGCTCCCTCGCGCGCGGTCCGCGCGCGCCCC 740
DB 554 CAGGTCCTCTGAGCAAAAGGACACCGCGCTCCCTCGCGCGCGGTCCGCGCGCGCCCC 613
QY 741 TGCCTGGAGCGCGCGGCGAGAGCTGGAGCCCTCTGCGCTACTGCGTCCCGTCAATCCAG 800
DB 614 TGCCTGGAGCGCGCGGCGAGAGCTGGAGCCCTCTGCGCTACTGCGTCCCGTCAATCCAG 673
QY 801 CGGACTCAGCCAGCGCGCGAGCTCCCGCCGAGAACGACACCGACCGAGCGGCTAC 860
DB 674 CGGACTCAGCCAGCGCGCGAGCTCCCGCCGAGAACGACACCGACCGAGCGGCTAC 733
QY 861 GCGCGGAGCGCGCGCGCGAGCGCGGAGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCG 920
DB 734 GCGCGGAGCGCGCGCGCGAGCGCGGAGCGCGGAGCGCGGAGCGCGCGCGCGCGCGCG 793

QY 921 GTACCATCAAGCAGCAGAGCTTCCGGGAGAGCTTCCGGCGGCCCGCCAGAGAGATGAAGCTG 980
DB 794 GTACCATCAAGCAGCAGAGCTTCCGGGAGAGCTTCCGGCGGCCCGCCAGAGAGATGAAGCTG 853
QY 981 GATTCCCGCGCGCGCGCAGCG 1040
DB 854 GATTCCCGCGCGCGCGCAGCG 913
QY 1041 GCGCTTCTGGGGCTG 1100
DB 914 GCGCTTCTGGGGCTG 973
QY 1101 CTGAGCTCGCTGCTGCGCTTCCGGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1160
DB 974 CTGAGCTCGCTGCTGCGCTTCCGGGAGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
QY 1161 GCTAC 1220
DB 1034 GCTAC 1093
QY 1221 GTGAGCGCGCTTCTGAGCAAGAGCGCGCTGAGAAAGTATCTGACCGCGCGCGCGCTGCC 1280
DB 1094 GTGAGCGCGCTTCTGAGCAAGAGCGCGCTGAGAAAGTATCTGACCGCGCGCGCGCTGCC 1153
QY 1281 GCG 1340
DB 1154 GCG 1213
QY 1341 GCT 1400
DB 1214 GCT 1273
QY 1401 CCGGAGAGCTGG 1460
DB 1274 CCGGAGAGCTGG 1333
QY 1461 GCG 1520
DB 1334 GCG 1393
QY 1521 CCGGAGAGCTGCC 1580
DB 1394 CCGGAGAGCTGCC 1453
QY 1581 TGAATCTTCTGCTGCTGCTGCC 1638
DB 1454 TGAATCTTCTGCTGCTGCTGCC 1511

RESULT 9

AAAL26629
ID AAL26629 standard; cDNA; 753 BP.

AC AAL26629;
XX

DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 19086.
DE

XX Human; breast cancer; cell marker; cytostatic; ss.
XX

OS Homo sapiens.
XX

XX WO200151628-A2.
XX

XX 19-JUL-2001.
PD

XX 10-JAN-2001; 2001WO-US000798.
PF

XX 14-JAN-2000; 2000US-0176077P.
XX

PR 14-MAR-2000; 2000US-0189167P.
PR

PR 24-MAR-2000; 2000US-0192099P.
PR

PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX New peptide useful as a marker for the diagnosis of breast cancer.
PT Claim 1; Page 3602; 3695pp; English.
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAU07544-AAU26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 753 BP; 236 A; 151 C; 138 G; 216 T; 0 U; 12 Other;
Query Match 17.3%; Score 629; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 8.4e-223;
Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2445 GATTCTATTATCATGATAGGATCTATACATACATCTCAAGTGCATCTATTC 2504
DB 62 GATTCTATTATCATGATAGGATCTATACATACATCTCAAGTGCATCTATTC 121
QY 2505 ACTCTCATTAATCCATCATGTTCTTAATTTTGTAACTTTACTGTAAATAAGTGCAC 2564
DB 122 ACTCTCATTAATCCATCATGTTCTTAATTTTGTAACTTTACTGTAAATAAGTGCAC 181
QY 2565 TGAATCTTCAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2624
DB 182 TGAATCTTCAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 241
QY 2625 CTATATCTTGTAAATTTCAAAAGTGAAGCAAGATTTAACTGGCCAGTTTGTATGCA 2684
DB 242 CTATATCTTGTAAATTTCAAAAGTGAAGCAAGATTTAACTGGCCAGTTTGTATGCA 301
QY 2685 AATGCTGTAAAGATATAGAAATGAAGTCTGTGAGGCTTCTTATCTCCAAGTCTATGTAT 2744
DB 302 AATGCTGTAAAGATATAGAAATGAAGTCTGTGAGGCTTCTTATCTCCAAGTCTATGTAT 361
QY 2745 TTCTGGAGACCAACACAGATACAGATTAATCACAAGAAAGCTTTTATTAAGGCTTA 2804
DB 362 TTCTGGAGACCAACACAGATACAGATTAATCACAAGAAAGCTTTTATTAAGGCTTA 421
QY 2805 AACCAAGACCTTGTCTAGATATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2864
DB 422 AACCAAGACCTTGTCTAGATATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
QY 2865 ACTTGATGTTATGTAAGGGGTGAGACACACAGCTCTGATATGATGAGGAAATATCT 2924
DB 482 ACTTGATGTTATGTAAGGGGTGAGACACACAGCTCTGATATGATGAGGAAATATCT 541
QY 2925 GGGTCTTTTCGTTCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2984
DB 542 GGGTCTTTTCGTTCAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
QY 2985 TGTGTTTAAACACCGTTTAAACTTTAGCCTTACAGGCTGCTTATGATAGTGTGTTA 3044
DB 602 TGTGTTTAAACACCGTTTAAACTTTAGCCTTACAGGCTGCTTATGATAGTGTGTTA 661
QY 3045 ATACATCCCAATTAATGATCTCTGACATGC 3073

DB 662 ATACATCCCAATTAATGATCTCTGACATGC 690
RESULT 10
ABX74439
ID ABX74439 standard; cDNA; 909 BP.
XX AC ABX74439;
XX 21-MAR-2003 (first entry)
XX Human cDNA sequence #11 up-regulated in CC-RCC patients.
XX Human; microarray; solid surface; immobilised probe; CC-RCC;
KW differential expression profile; aggressive CC-RCC tumour type;
KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW gene expression profiling; tumour tissue; gene; ss.
XX Homo sapiens.
OS WO200279411-A2.
PN 10-OCT-2002.
XX 29-MAR-2002; 2002WO-US009576.
XX 29-MAR-2001; 2001US-0279411P.
XX (VAND-) VAN ANDEL INST.
PI Haab B, Rhodes D, Teh BT, Takashi M;
XX WPI; 2003-040679/03.
XX New microarray, comprising a matrix of cDNA probe from a set of probes
PT immobilized to a solid surface in predetermined order, useful in the
PT prognosis of patients with clear cell renal carcinoma.
XX Claim 35; Page 119-120; 179pp; English.
XX The present invention relates to a microarray comprising a matrix of at
CC least one cDNA probe from a set of probes immobilised to a solid surface
CC in a predetermined order, where a row of pixels corresponds to replicates
CC of one distinct probe from the set. The probes are complementary to
CC nucleic acid sequences that are expressed differentially in aggressive as
CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
CC and that hybridise to the probes under high stringency conditions. The
CC microarray is useful for the prognosis of patients with CC-RCC, wherein
CC aggressive and non-aggressive CC-RCC tumour types are characterised by
CC differential expression profiles of genes that hybridise with one or more
CC probes immobilised on the microarray. The arrays are useful for gene
CC expression profiling of tumour and normal tissues. The present sequence
CC represents a human cDNA sequence up-regulated in CC-RCC patients
XX
SQ Sequence 909 BP; 218 A; 292 C; 256 G; 143 T; 0 U; 0 Other;
Query Match 15.2%; Score 552; DB 10; Length 909;
Best Local Similarity 100.0%; Pred. No. 2.5e-194;
Matches 552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 369 ACTCTGGGACATCTGAGAAAGCTGTAGTCTTGGATTAATCTTGAACACCTTAAAGCT 428
DB 181 ACTCTGGGACATCTGAGAAAGCTGTAGTCTTGGATTAATCTTGAACACCTTAAAGCT 240
QY 429 TTAACCGCCTTAACCGGACCAAGCATCAGAGATAATTTGCTTTACAGAAATGGGAGCGA 488
DB 241 TTAACCGCCTTAACCGGACCAAGCATCAGAGATAATTTGCTTTACAGAAATGGGAGCGA 300
QY 489 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCTTCCACTCGGATTTCAACATCC 548
DB 301 TCTCTGAAATCGCCCATTCAGTCCGACTTGGATGCTTCCACTCGGATTTCAACATCC 360
QY 549 GCCAAAGAGTCTTTCGAATACCTCTCCCGTTTGTAGAGCTGGACACCCAGGGAGCGCG 608

```
Db 361 GCCAAGAAGCTCTTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGAGCGCGG 420
QY 609 TGTGTCAGCTGATCAACACATTCGACGCGGTGGCCACCCAGTCTTTGTCCTCCACCGCGAG 668
Db 421 TGTGTCAGCTGATCAACACATTCGACGCGGTGGCCACCCAGTCTTTGTCCTCCACCGCGAG 480
QY 669 CTGTTGACTCAACAGAGTCCCTCTGAGCAAGAGGACCGCGCTCCCTCGGCGCGCGGTC 728
Db 481 CTGTTGACTCAACAGAGTCCCTCTGAGCAAGAGGACCGCGCTCCCTCGGCGCGCGGTC 540
QY 729 GCGGCGCGCGCGCTGCTGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 788
Db 541 GCGGCGCGCGCGCTGCTGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 789 CCGGTCAATCCAGCGGACTCAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 848
Db 601 CCGGTCAATCCAGCGGACTCAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 849 GACAGCGGCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 908
Db 661 GACAGCGGCTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 909 GGGGCGAGCGCG 920
Db 721 GGGGCGAGCGCG 732
```

RESULT 11

```
ID ADQ24375 standard; DNA; 1573 BP.
XX ADQ24375;
XX AC
XX ADQ24375;
XX DT
XX 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 7195.
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.
XX Homo sapiens.
XX OS
XX WO2004048938-A2.
XX PN
XX 10-JUN-2004.
XX PD
XX 26-NOV-2003; 2003WO-US038193.
XX PF
XX 26-NOV-2002; 2002US-0429739P.
XX PR
XX (PROT-) PROTEIN DESIGN LABS INC.
XX PA
XX Aziz N, Ginsburg WM, Zlotnik A;
XX PI
XX WPI; 2004-441208/41.
XX DR
XX
```

Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

Example 2; SEQ ID NO 7195; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic

acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the CC specification per se but was submitted in CD format by the inventor.

Sequence 1573 BP; 508 A; 312 C; 313 G; 411 T; 0 U; 29 Other;

Query Match 14.2%; Score 516; DB 12; Length 1573;
Best Local Similarity 99.8%; Pred. No. 4.9e-181;
Matches 636; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 2758 AACAGATACACAGATAATCAAGAAAGCTTTTAAAGGCTTAAACAGACCTTG 2817
Db 251 AACAGATACACAGATAATCAAGAAAGCTTTTAAAGGCTTAAACAGACCTTG 310
QY 2818 TCTAGATATTTTAGTTTGGTCCAGGTAGCAGTGTGAGAAATCTCAGTGGATGTAT 2877
Db 311 TCTAGATATTTTAGTTTGGTCCAGGTAGCAGTGTGAGAAATCTCAGTGGATGTAT 370
QY 2878 GTAAGGGGTGAGACACACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTCGTC 2937
Db 371 GTAAGGGGTGAGACACACAGTCTGACTATGAGTGAGGAAAAATATCTGGTCTTCGTC 430
QY 2938 AGTTTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2997
Db 431 AGTTTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
QY 2998 CGTTAAACTCTTAGCCTACAGGTGGCTCTTATGTACATAGTTGTTAATACATCCCAATTA 3057
Db 491 CGTTAAACTCTTAGCCTACAGGTGGCTCTTATGTACATAGTTGTTAATACATCCCAATTA 550
QY 3058 ATGATGTCTGACATGCTATTTTGTAGGGAGAAAAATATGCTAAATGATATTTTGAGTTA 3117
Db 551 ATGATGTCTGACATGCTATTTTGTAGGGAGAAAAATATGCTAAATGATATTTTGAGTTA 610
QY 3118 AAATATCTTTTGGGAGGATTTGCTGAAAGTTGCACCTTTTGTACATGCTTATGCTTG 3177
Db 611 AAATATCTTTTGGGAGGATTTGCTGAAAGTTGCACCTTTTGTACATGCTTATGCTTG 670
QY 3178 GTACAAGCTTATGCTGCTTAAATATTTTAAAAAATAAATACTGCTGTCGAGAAACC 3237
Db 671 GTACAAGCTTATGCTGCTTAAATATTTTAAAAAATAAATACTGCTGTCGAGAAACC 729
QY 3238 AGCTGGTTTGAAGAAAGTTTAGTAGTGCATGACGATAAATAGATAATCTTATTTAGT 3297
Db 730 AGCTGGTTTGAAGAAAGTTTAGTAGTGCATGACGATAAATAGATAATCTTATTTAGT 789
QY 3298 ATTTTCAGCACTCCATAAATTTTATTACCTAAATATTTGCCACACTATTTTGTGATTAAA 3357
Db 790 ATTTTCAGCACTCCATAAATTTTATTACCTAAATATTTGCCACACTATTTTGTGATTAAA 849
QY 3358 AATTCTTACTAAGGAATAAAACTTTTAATATACGATA 3394
Db 850 AATTCTTACTAAGGAATAAAACTTTTAATATACGATA 886
```

RESULT 12

AAS26175
ID AAS26175 standard; cDNA; 498 BP.

XX AAS26175;

XX 07-NOV-2001 (first entry)

Human cDNA encoding a novel secreted protein, Seq ID 354.

Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; vulnary; secreted protein; rheumatoid arthritis; hyperproliferative disorder; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder; corneal infection;

wound healing; epithelial cell proliferation; skin ageing; food additive;
preservative; antiproliferative.

Homo sapiens.
WO200155322-A2.

25-SEP-2000; 2000US-02344997P.
25-SEP-2000; 2000US-02344998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
13-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239355P.
13-OCT-2000; 2000US-0239377P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249246P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256179P.
06-DEC-2000; 2000US-0256179P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251859P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.

PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR P-PSDB; ABU55256.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 1; SEQ ID NO 354; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
XX Sequence 498 BP; 142 A; 137 C; 115 G; 104 T; 0 U; 0 Other;
SQ

Query Match 13.6%; Score 496; DB 8; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.4e-173;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 241 TGAACGAGCAGCACCACGAGATACCTTCAAAATACCGCACAGATTAATAGAAAGAAA 300
DB 1 TGAACGAGCAGCACCACGAGATACCTTCAAAATACCGCACAGATTAATAGAAAGAAA 60
QY 301 GAAGAGACCGCAATTAATGAATGCTTGTCTGAGCTGAAAGATTACTGCTGCAACATCTGA 360
DB 61 GAAGAGACCGCAATTAATGAATGCTTGTCTGAGCTGAAAGATTACTGCTGCAACATCTGA 120
QY 361 AATTGACAACTCTGGGACATCTGAGAAAGCTGTAGTCTTGGATTAACCTTGAACACT 420
DB 121 AATTGACAACTCTGGGACATCTGAGAAAGCTGTAGTCTTGGATTAACCTTGAACACT 180
QY 421 TAAAGCTTTTAAACCGCTTTAAACCGCAGCAACAGCATCAGAGATTAATGCTTTACAGATG 480
DB 181 TAAAGCTTTTAAACCGCTTTAAACCGCAGCAACAGCATCAGAGATTAATGCTTTACAGATG 240
QY 481 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTGGATTCGCTCCACTCGGGATTC 540
DB 241 GGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTGGATTCGCTCCACTCGGGATTC 300
QY 541 AAAATGCGGCAAGAAAGCTTTTGAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGG 600
DB 301 AAAATGCGGCAAGAAAGCTTTTGAATACCTCTCCCGTTTGAGAGCTGGACACCCAGGG 360

QY 601 AGCCGCGGTGTCTCAGCTGATCAACCACTTTCACCGCGTGGCCACCCAGTTCTTGCCCA 660
DB 361 AGCCGCGGTGTCTCAGCTGATCAACCACTTTCACCGCGTGGCCACCCAGTTCTTGCCCA 420
QY 661 CCCCAGCTGTGTACTCAACAGGTCCTCTGAGCAAAAGGACCGGCGCTCCCTCGGCGG 720
DB 421 CCCCAGCTGTGTACTCAACAGGTCCTCTGAGCAAAAGGACCGGCGCTCCCTCGGCGG 480
QY 721 CCGGGTCCGCGCGCGC 736
DB 481 CCGGGTCCGCGCGCGC 496

RESULT 14

ACC72415
ID ACC72415 standard; cDNA; 430 BP.
XX
XX AC ACC72415;
XX
XX DT 07-JUL-2003 (first entry)
XX Human cDNA insert from ovarian cancer clone O1668P.
XX
XX DE Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer; gene;
XX KW ss.
XX OS Homo sapiens.
XX PN WO2003029468-A1.
XX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031467.
XX PR 02-OCT-2001; 2001US-0327135P.
XX PR 30-MAY-2002; 2002US-0384531P.
XX (CORI-) CORIXA CORP.
XX PA Algate PA, Mannion J;
XX PI WPI; 2003-372001/35.
XX
XX PT New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.
PS Claim 1; Page 147; 169pp; English.
XX
XX The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention
XX
XX Sequence 430 BP; 162 A; 81 C; 61 G; 126 T; 0 U; 0 Other;
SQ

Query Match 11.8%; Score 430; DB 9; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.9e-149;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2240 AAATGGTATCTCTTAGTAACCTTGCCTGTTAAAGAAACACGGAGCTGGGCATGTCAG 2299
DB 1 AAATGGTATCTCTTAGTAACCTTGCCTGTTAAAGAAACACGGAGCTGGGCATGTCAG 60
QY 2300 AACTAAGTCAGGAGGAGATGGATGAGAGCCAGATCATTCCTAGTACATTGTCTAA 2359
DB 61 AACTAAGTCAGGAGGAGATGGATGAGAGCCAGATCATTCCTAGTACATTGTCTAA 120
QY 2360 CACTTTATTGAGAAATTCACCATGAATTAATGGACTCATCTTAATTTCTTAAAGTCCAT 2419
DB 121 CACTTTATTGAGAAATTCACCATGAATTAATGGACTCATCTTAATTTCTTAAAGTCCAT 180

Qy	2420	ATATAGATAGATATCTATCTGTACAGATTTCTATTATTCATAGATAGGTATCTATACAT	2479
Db	181	ATATAGATAGATATCTATCTGTACAGATTTCTATTATTCATAGATAGGTATCTATACAT	240
Qy	2480	ACATCTCTCAAGTGCATCTATTCCCACTCTCATTAATCAATCATGTTCCCTAAATTTTTGT	2539
Db	241	ACATCTCTCAAGTGCATCTATTCCCACTCTCATTAATCAATCATGTTCCCTAAATTTTTGT	300
Qy	2540	AACTTACTGTAAAAAAGTGCACTGAACTTCAAAAACAAAAACAAACAACAACA	2599
Db	301	AACTTACTGTAAAAAAGTGCACTGAACTTCAAAAACAAAAACAAACAACAACA	360
Qy	2600	AAAAACAAGTCCAAACATGATATATCCCTATATCTGTTAAATTCAAAAGTGAACGAAGC	2659
Db	361	AAAAACAAGTCCAAACATGATATATCCCTATATCTGTTAAATTCAAAAGTGAACGAAGC	420
Qy	2660	ATTTAACTGG	2669
Db	421	ATTTAACTGG	430
RESULT 15			
ACD19339			
ID	ACD19339 standard; cDNA; 628 BP.		
XX	AC	ACD19339;	
XX	AC		
DT	25-AUG-2003 (first entry)		
XX			
DE	cDNA encoding novel human protein #19.		
XX			
KW	Human; NOV; gene therapy; endocrine related disease; diabetes;		
KW	metabolism-related disease; obesity; central nervous system disorder;		
KW	Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;		
KW	schizophrenia; depression; autoimmune disorder; inflammatory disorder;		
KW	psoriasis; allergy; lupus erythematosus; asthma; cancer;		
KW	inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;		
KW	colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;		
KW	prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;		
KW	lung disease; emphysema; obstructive pulmonary disease; haemophilia;		
KW	stroke; infection; gene; ss.		
OS	Homo sapiens.		
XX			
PN	WO2003023002-A2.		
XX			
PD	20-MAR-2003.		
XX			
PF	03-SEP-2002; 2002WO-US028539.		
XX			
PR	07-SEP-2001; 2001US-0318120P.		
PR	10-SEP-2001; 2001US-0318130P.		
PR	10-SEP-2001; 2001US-0318430P.		
PR	17-SEP-2001; 2001US-0322636P.		
PR	17-SEP-2001; 2001US-0322781P.		
PR	17-SEP-2001; 2001US-0322816P.		
PR	17-SEP-2001; 2001US-0322817P.		
PR	19-SEP-2001; 2001US-0323519P.		
PR	20-SEP-2001; 2001US-0323631P.		
PR	20-SEP-2001; 2001US-0323636P.		
PR	23-SEP-2001; 2001US-0324969P.		
PR	23-SEP-2001; 2001US-0325091P.		
PR	26-SEP-2001; 2001US-0324990P.		
PR	17-APR-2002; 2002US-0373212P.		
PR	06-SEP-2002; 2002US-00236177.		

PI	Lapley DM, Edinger SR, Burgess CE;
XX	
DR	WFI; 2003-313242/30.
DR	P-PSDB; ABO14646.
XX	
PT	New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
FT	and polynucleotides, useful in gene therapy, e.g. for treating or
PT	preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT	stroke or infections.
XX	
FS	Claim 20; Page 125-126; 586pp; English.
XX	
CC	The invention describes a new isolated polypeptide (NOVX). The NOVX
CC	polypeptide, nucleic acid and antibody are useful as therapeutics,
CC	particularly in the manufacture of a medicament for treating a syndrome
CC	associated with a human disease, which includes a pathology associated
CC	with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC	therapy for treating the disease or condition. In particular, the NOVX
CC	polypeptide or polynucleotide is useful for treating endocrine/ metabolism-related diseases (e.g. obesity or diabetes), central nervous system disorders (e.g. Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus, asthma, inflammatory bowel disease, rheumatoid arthritis or osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian, prostate or brain cancers, or melanoma), liver diseases (e.g. liver cirrhosis), lung diseases (emphysema or obstructive pulmonary disease), haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC	These are also useful in developing powerful assay system for functional
CC	analysis of various human disorders, as well as in diagnostic
CC	applications, and for monitoring the effects of drugs during clinical
CC	trials. This sequence encodes a novel human NOV protein
XX	
SQ	Sequence 628 BP; 189 A; 157 G; 147 G; 135 T; 0 U; 0 Other;
	Query Match 11.7%; Score 427; DB 10; Length 628;
	Best Local Similarity 99.8%; Pred.No. 4.7e-148;
	Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	176 GGAAACATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTGAACCCAAAAG 235
Db	48 GGAAACATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTGAACCCAAAAG 107
QY	236 GAGCNTGAACGAGACGACACCAGAAGATCCTTACAANTTACCGCACAGATTATAGAAA 295
Db	108 GAGCATGAACGAGACGACACCAGAAGATCCTTACAANTTACCGCACAGATTATAGAAA 167
QY	296 GAAAAAAGAGACCGCAANTTAATGAATGCATTGCTCAGCTGAAAGATTTCCTGCTCTGAACA 355
Db	168 GAAAAAAGAGACCGCAANTTAATGAATGCATTGCTCAGCTGAAAGATTTCCTGCTCTGAACA 227
QY	356 TCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTAGTCTTGGAAATTAACCTTGAA 415
Db	228 TCTGAAATTGACAACTCTGGGACATCTGGAGAAAGCTAGTCTTGGAAATTAACCTTGAA 287
QY	416 ACACCTTAAAAGCTTTTAAACCGCCCTTAAACCGAGCAACAGCATCAGAAGATAATTCCTTTACA 475
Db	288 ACACCTTAAAAGCTTTTAAACCGCCCTTAAACCGAGCAACAGCATCAGAAGATAATTCCTTTACA 347
QY	476 GAATGGGAGCGATCTCTGAANTCGGCCAATTAGTCGCAATTGGATCGGTTCCACTCGGG 535
Db	348 GAATGGGAGCGATCTCTGAANTCGGCCAATTAGTCGCAATTGGATCGGTTCCACTCGGG 407
QY	536 ATTTTCAAACATGGGCCCCAAGAAGCTCTTGAATACCTCTCCCGGTTTTGAGAGCTGGACACC 595
Db	408 ATTTTCAAACATGGGCCCCAAGAAGCTCTTGAATACCTCTCCCGGTTTTGAGAGCTGGACACC 467
QY	596 CAGGGAGCCCGGTGTGTCCAGCTGATCAACCACTTGCACGCCGTGGCACCCAGTTC 653
Db	468 CAGGGAGCCCGGTGTGTCCAGCTGATCAACCACTTGCACGCCGTGGCACCCAGTTC 525

Search completed: December 17, 2004, 05:53:25

Mon Dec 20 08:49:05 2004

us-10-078-650-1.Oligo.rng

Page 22

Job time : 1668.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 03:15:04 ; Search time 15113.1 Seconds
(without alignments)
11392.870 Million cell updates/sec

Title: US-10-078-650-1
Perfect score: 3641
Sequence: 1 cgcgactgaaggagagagc.....ttaaaaaaaaaaaaaaaaaa 3641

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_btg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3641	100.0	3641	6	BD011858	BD011858 Novel BHL
2	3641	100.0	3641	9	AB044088	AB044088 Homo sapi
3	3043	83.6	3431	9	BC025968	BC025968 Homo sapi
4	2853	78.4	204228	9	AC022509	AC022509 Homo sapi
5	2290	62.9	3274	6	BD176928	BD176928 Novel BHL
6	1583	43.5	2304	6	AX591648	AX591648 Sequence
7	1378	37.8	1511	6	BD011867	BD011867 Novel BHL
8	921	25.3	1973	9	AK094422	AK094422 Homo sapi
9	629	17.3	753	6	CQ434066	CQ434066 Sequence
10	552	15.2	909	6	CQ727066	CQ727066 Sequence
11	450	12.6	593	9	HS0270695	HS0270695 Homo sapi
12	413	11.3	797	11	BV208394	BV208394 BHLHB3_20
13	381	10.7	462	6	CQ424141	CQ424141 Sequence
14	386	10.6	447	6	CQ432989	CQ432989 Sequence
15	353	10.0	471	6	CQ685084	CQ685084 Sequence
16	335	9.8	521	6	AX198436	AX198436 Sequence
17	354	9.7	501	6	AX208985	AX208985 Sequence
18	354	9.7	501	6	AX209655	AX209655 Sequence
19	341	9.4	580	6	AX198458	AX198458 Sequence

20	341	9.4	580	6	AX209005	AX209005 Sequence
21	339	9.3	411	6	CQ415230	CQ415230 Sequence
22	312	8.6	497	6	CQ680194	CQ680194 Sequence
23	308	8.5	553	6	CQ410788	CQ410788 Sequence
24	308	8.5	684	6	CQ398115	CQ398115 Sequence
25	308	8.5	684	6	CQ404404	CQ404404 Sequence
26	296	8.1	409	6	AX209663	AX209663 Sequence
27	296	8.1	502	6	AX198354	AX198354 Sequence
28	296	8.1	502	6	AX208910	AX208910 Sequence
29	281	7.7	400	6	CQ712682	CQ712682 Sequence
30	274	7.5	285	6	AX209128	AX209128 Sequence
31	274	7.5	397	6	AX198590	AX198590 Sequence
32	271	7.4	322	6	AX208788	AX208788 Sequence
33	271	7.4	424	6	AX198219	AX198219 Sequence
34	266	7.3	440	6	AX341473	AX341473 Sequence
35	265	7.3	404	6	CQ682847	CQ682847 Sequence
36	252	6.9	278	6	CQ695845	CQ695845 Sequence
37	246	6.8	502	6	AX8932348	AX8932348 Sequence
38	246	6.8	502	6	BD027881	BD027881 Sequence
39	242	6.6	604	6	CQ412716	CQ412716 Sequence
40	240	6.6	364	6	CQ697079	CQ697079 Sequence
41	239	6.6	480	6	CQ685093	CQ685093 Sequence
42	237	6.5	261	6	CQ687963	CQ687963 Sequence
43	234	6.4	260	6	CQ695847	CQ695847 Sequence
44	226	6.2	260	6	CQ688754	CQ688754 Sequence
45	225	6.2	265	6	CQ686173	CQ686173 Sequence

ALIGNMENTS

RESULT 1
BD011858
LOCUS BD011858 3641 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel BHLH type transcriptional gene, DEC2.
ACCESSION BD011858
VERSION BD011858.1 GI:22092047
KEYWORDS WO 0114551-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3641)
AUTHORS Fujimoto,K., Shin,M. and Kato,Y.
TITLE Novel BHLH type transcriptional gene, DEC2
JOURNAL Patent: WO 0114551-A 1 01-MAR-2001.
COMMENT CHUGAI PHARM CO LTD,KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
OS Homo sapiens (human)
PN WO 0114551-A/1
PD 01-MAR-2001
PF 19-JUN-2000 WO 2000JP003991
PR 19-AUG-1999 JP 99p 233286
PI KATSUMI FUJIMOTO,MEI SHIN,YUKIO KATO
PC C12N15/12,C07L14/475,C07K16/18
CC
FH Key Location/Qualifiers
FT CDS (135)..(1580).
FEATURES
source
1..3641
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 3641; DB 6; Length 3641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCACCTGAAGAGGAGCGAGAGAGACTGGAGACGACAGATCCCCCAAGGTC 60
Db 1 CTGCACCTGAAGAGGAGCGAGAGAGACTGGAGACGACAGATCCCCCAAGGTC 60
Qy 61 CCAAGCTACCGTCCCGACAGATATTATTACAGAGCCCCCAAAATCGAAGACAGAGAAACG 120

61	Db		CCAAGCTACCGTCCCA	CAGAGTAA	TTGTGTACAGAGCCCAAAA	TGAAACAGAGGAAACG	120
121	QY		AACAGCAGTTGAACAT	GTGACCAAGGAAT	TCTCATTTTGGCAAGAGACAGAGT	TACTTGGAAAC	180
121	Db		AACAGCAGTTGAA	CAATGAGCGAGGAAT	TCTCATTTGCAAGAGACAGAGT	TACTTGGAAAC	180
181	QY		A7AGAGATTTATAGCA	TGCACTATT	CTCTTTTGTATATGTGTATAA	CCCAAAAGGACGA	240
181	Db		ATAGAGATTTTATAGCA	CTGCACTATT	CCCTTTTGTATATGTGTATAA	CCCAAAAGGACGA	240
241	QY		TGAAACGAGACACAC	CAAGGATACCTTACAAAT	TACCGCACAGATTAATAGAAAGAAA	300	
241	Db		TGAAACGAGACACAC	CAAGGATACCTTACAAAT	TACCGCACAGATTAATAGAAAGAAA	300	
301	QY		GAAGAGACCGAATTAAT	GAATGCATTCTCAGCTGAAAGATTTACT	TGCGCTGAACATCTGA	360	
301	Db		GAAGAGACCGAATTAAT	GAATGCATTCTCAGCTGAAAGATTTACT	TGCGCTGAACATCTGA	360	
361	QY		AATTGACAACTCTGGGA	CATCTGGAGAAAGCTGTAGTCTTTGGAATTAAC	TTTGAACACT	420	
361	Db		AATTGACAACTCTGGGA	CATCTGGAGAAAGCTGTAGTCTTTGGAATTAAC	TTTGAACACT	420	
421	QY		TAAAGCTTTAAACCGCTT	AAACCGAGCAACAGATCAGAAGATAA	TTGCTTTTACAGAATG	480	
421	Db		TAAAGCTTTAAACCGCTT	AAACCGAGCAACAGATCAGAAGATAA	TTGCTTTTACAGAATG	480	
481	QY		GGGAGCGATCTCTGAAAT	TGCGCCATTCAGTCCGACTTGGATGGTTC	CACTCGGATTTTC	540	
481	Db		GGGAGCGATCTCTGAAAT	TGCGCCATTCAGTCCGACTTGGATGGTTC	CACTCGGATTTTC	540	
541	QY		AAACATATGCGCCAAAGAA	GCTTTGCAATACCTCTCCGGTTTGAGAGCTCGACAC	CCACGAGG	600	
541	Db		AAACATATGCGCCAAAGAA	GCTTTGCAATACCTCTCCGGTTTGAGAGCTCGACAC	CCACGAGG	600	
601	QY		AGCCGCGTGTGTCCAGT	TGATCAACCACTTGCAACCGCTGGCCACCCAGTTCTTTGCCCA	660		
601	Db		AGCCGCGTGTGTCCAGT	TGATCAACCACTTGCAACCGCTGGCCACCCAGTTCTTTGCCCA	660		
661	QY		CCCCGACGTGTGTGACT	CAACAGCTCCCTCTGAGCAAAAGCAACCGCGCTCCCTCGGCGG	720		
661	Db		CCCCGACGTGTGTGACT	CAACAGCTCCCTCTGAGCAAAAGCAACCGCGCTCCCTCGGCGG	720		
721	QY		CCGGGTCCGCGGCCCGCCCT	CTCGAGGCGCGGCGCAGAGCTGAGCGCCCTCGGCT	780		
721	Db		CCGGGTCCGCGGCCCGCCCT	CTCGAGGCGCGGCGCAGAGCTGAGCGCCCTCGGCT	780		
781	QY		ACTGTGTGCGGTGAT	CCAGCGGACTCAGCGCAAGCGCGGCTCGCGCCGAGAGACGA	840		
781	Db		ACTGTGTGCGGTGAT	CCAGCGGACTCAGCGCAAGCGCGGCTCGCGCCGAGAGACGA	840		
841	QY		CGGACACCGACACAGCGGCT	TACGGCGCGCAAGCCGAGCCCGCGACCCGAGAAAGCA	900		
841	Db		CGGACACCGACACAGCGGCT	TACGGCGCGCAAGCCGAGCCCGCGACCCGAGAAAGCA	900		
901	QY		AAGCGCGGCGGCGAGCGCGCT	CAACATCAAGCAGAGAGCTCCCGGGGAGAGACTCGCGG	960		
901	Db		AAGCGCGGCGGCGAGCGCGCT	CAACATCAAGCAGAGAGCTCCCGGGGAGAGACTCGCGG	960		
961	QY		CGCCCAAGAGGATGAAGCT	TGATTTCCCGCGCGCGCGCAGCGCGCGCGCGCGGCGG	1020		
961	Db		CGCCCAAGAGGATGAAGCT	TGATTTCCCGCGCGCGCGCAGCGCGCGCGCGCGGCGG	1020		
1021	QY		GGCGCGCGCGCGCGCAGCGCGT	TTCTGGGCGCGACCGCTGCGCGCGCGCGCGCTGCG	1080		
1021	Db		GGCGCGCGCGCGCGCAGCGCGT	TTCTGGGCGCGCGACCGCTGCGCGCGCGCGCGCTGCG	1080		
1081	QY		TCAGACCCGACGCGCGCT	GTAGCTCGCTGTGTGGCTTCGCGGAGCGGAGCGCGCG	1140		
1081	Db		TCAGACCCGACGCGCGCT	GTAGCTCGCTGTGTGGCTTCGCGGAGCGGAGCGCGCG	1140		
1141	QY		CTTCTCCGCGCGCGCGCGCGCCCT	TTCTGCTGCGCTTCTGCTTCTCTCTCGC	1200		

Db	1141	CCTTCCGCGAGCCGGGGCCGCGCGGCCCTTCTGCTCGCCCTTCGTGCTTCCCTCTCGC	1200
Qy	1201	CTTCTGCAGCTGCCGCTACGTGCAGCCCTTCTCTGCAAGAGCGCCTTGAGAAGTATC	1260
Db	1201	CTTCTGCAGCTGCCGCTACGTGCAGCCCTTCTGCAAGAGCGCCTTGAGAAGTATC	1260
Qy	1261	TGTACCCGGGGGGGTGCCGCCCGGTTCCGCTGCTATACCCCGGATCCCCCGCCCCG	1320
Db	1261	TGTACCCGGGGGGGTGCCGCCCGGTTCCGCTGCTATACCCCGGATCCCCCGCCCCG	1320
Qy	1321	CGGAGCGCGGCGAGCGCCCGCGCGCTGCCCGCGCGCGCGGCTTCCCTGCTGCTGT	1380
Db	1321	CGGAGCGCGGCGAGCGCCCGCGCGCTGCCCGCGCGCGCGGCTTCCCTGCTGCTGT	1380
Qy	1381	CTTCCGTTGTCGCCCTTCCGAGAGCGGGCGCGCGCGGACCTCTCTGCGCG	1440
Db	1381	CTTCCGTTGTCGCCCTTCCGAGAGCGGGCGCGCGCGGACCTCTCTGCGCG	1440
Qy	1441	ACGAGTGGCGCCCTTTGGGGCGCGCACCCCGACACCGCAGCGCGCAGCCACCTGC	1500
Db	1441	ACGAGTGGCGCCCTTTGGGGCGCGCACCCCGACACCGCAGCGCGCAGCCACCTGC	1500
Qy	1501	CCTTCCGCGGCCCGCGAGCCCGGGAAACCGGAGAGCTTGCTCAGGAAGATCCCTCGC	1560
Db	1501	CCTTCCGCGGCCCGCGAGCCCGGGAAACCGGAGAGCTTGCTCAGGAAGATCCCTCGC	1560
Qy	1561	AGCAGGAAGAGATCCTCGTAATCTTTCGCTCCCGAGGACGGAGGTTCAAGCAGAGT	1620
Db	1561	AGCAGGAAGAGATCCTCGTAATCTTTCGCTCCCGAGGACGGAGGTTCAAGCAGAGT	1620
Qy	1621	GAGAAGTAAAAATACCCCTTAAGGAGGTTCAAGCAGAGTCAGAAATAAAAATACCTTAAG	1680
Db	1621	GAGAAGTAAAAATACCCCTTAAGGAGGTTCAAGCAGAGTCAGAAATAAAAATACCTTAAG	1680
Qy	1681	GTCTTTAAGGGAGAGTGTAATAGATGCAGCAGCAGGCATAAACAGAAACAACAACAG	1740
Db	1681	GTCTTTAAGGGAGAGTGTAATAGATGCAGCAGCAGGCATAAACAGAAACAACAACAG	1740
Qy	1741	GTGTTATGTGCATATCGGAGTTCCTGTTTGTCTATCCGCGACACCCACCCCTCCACA	1800
Db	1741	GTGTTATGTGCATATCGGAGTTCCTGTTTGTCTATCCGCGACACCCACCCCTCCACA	1800
Qy	1801	CACATAACCTCTTCTTCCCCCAACAGCTGTAAAGATCTATGCGAAGACACTGGC	1860
Db	1801	CACATAACCTCTTCTTCCCCCAACAGCTGTAAAGATCTATGCGAAGACACTGGC	1860
Qy	1861	TCCTTTTAAATCCCCCAATAAATTTTGCCCCCTTTTAGGGCATGTTCCATTATCTCT	1920
Db	1861	TCCTTTTAAATCCCCCAATAAATTTTGCCCCCTTTTAGGGCATGTTCCATTATCTCT	1920
Qy	1921	TAAAAATGGAACCTAATTCGAGAGGAGTAAGAGGCTGTGTTCTGTGGCTGAGCTAGGT	1980
Db	1921	TAAAAATGGAACCTAATTCGAGAGGAGTAAGAGGCTGTGTTCTGTGGCTGAGCTAGGT	1980
Qy	1981	GAACCCGGGTAGGGGAAGATGTTAAACACTTTCAGCGCTTTTGAGTTGACATGGAAC	2040
Db	1981	GAACCCGGGTAGGGGAAGATGTTAAACACTTTCAGCGCTTTTGAGTTGACATGGAAC	2040
Qy	2041	AGCAGGTAGTGTATGTAGAGCTAGTTCTCAAAAGCTGCCCTGCCTTTTAGGAGGCGT	2100
Db	2041	AGCAGGTAGTGTATGTAGAGCTAGTTCTCAAAAGCTGCCCTGCCTTTTAGGAGGCGT	2100
Qy	2101	TCACAAAACAGATTGAGGCTCTTTTAGAATTTGAATTTACTCTTCAGTATTTCTAATGT	2160
Db	2101	TCACAAAACAGATTGAGGCTCTTTTAGAATTTGAATTTACTCTTCAGTATTTCTAATGT	2160
Qy	2161	TCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGGATCGAGTTCTCACGTTGCA	2220
Db	2161	TCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGGATCGAGTTCTCACGTTGCA	2220
Qy	2221	ACCTATCTGAAGTGGTTAAATGGTATCTCTTAGTAACTTTGCACCTCGTTAAGAACAC	2280
Db	2221	ACCTATCTGAAGTGGTTAAATGGTATCTCTTAGTAACTTTGCACCTCGTTAAGAACAC	2280

```
QY 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGATGAGAAAGCCAGCAATCA 2340
DB 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGGAAGGAGATGATGAGAAAGCCAGCAATCA 2340
QY 2341 TTCCTAGTACATTTGCTACACCTTTATTCAGAAATTCACCATGAATTAATGACATCATCT 2400
DB 2341 TTCCTAGTACATTTGCTACACCTTTATTCAGAAATTCACCATGAATTAATGACATCATCT 2400
QY 2401 TAAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
DB 2401 TAAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTTATCCA 2460
QY 2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTTCCCACTCTCATTAATCCAT 2520
DB 2461 TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTTCCCACTCTCATTAATCCAT 2520
QY 2521 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAATAAAGTGCACCTTCAAAACAAA 2580
DB 2521 CATGTTCCCTAAATTTTGTAAATCTTACTGTAAATAAAGTGCACCTTCAAAACAAA 2580
QY 2581 ACAAAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 2640
DB 2581 ACAAAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 2640
QY 2641 TTCAAAAGTGAAGCAAGCATTTAACTGSCCAGTTTGTGATTCGAAATGCTGTAAAGATAT 2700
DB 2641 TTCAAAAGTGAAGCAAGCATTTAACTGSCCAGTTTGTGATTCGAAATGCTGTAAAGATAT 2700
QY 2701 AGAATGAATCTCTGTGAGCCCTTCTATCTCCAACTATATGCTATGCTATGCTGAGACCAAC 2760
DB 2701 AGAATGAATCTCTGTGAGCCCTTCTATCTCCAACTATATGCTATGCTGAGACCAAC 2760
QY 2761 CAGATACACAGATAACACAAAGAAAGCTTTTATTAAGGCTTAACCAAGACCTTGCT 2820
DB 2761 CAGATACACAGATAACACAAAGAAAGCTTTTATTAAGGCTTAACCAAGACCTTGCT 2820
QY 2821 AGATATTTTTAGTTTGTGCCAAGTAGCAGCTGAGAAATCTCACTTGGATGTTATGTA 2880
DB 2821 AGATATTTTTAGTTTGTGCCAAGTAGCAGCTGAGAAATCTCACTTGGATGTTATGTA 2880
QY 2881 AGGGGTGACACACACAGCTCTGACATGATGAGTGAAGAAATATCTGGGCTTTTCTGCTAGT 2940
DB 2881 AGGGGTGACACACACAGCTCTGACATGATGAGTGAAGAAATATCTGGGCTTTTCTGCTAGT 2940
QY 2941 TTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
DB 2941 TTGGTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
QY 3001 TAAACTTTAGCCTACAGGTGGCTCTTATGACATAGTTGTTAATACATCAATTAATG 3060
DB 3001 TAAACTTTAGCCTACAGGTGGCTCTTATGACATAGTTGTTAATACATCAATTAATG 3060
QY 3061 ATGCTGACATGCTATTTTGTAGGAGAAATATGCTGTAATGATATTTTGAGTTAAA 3120
DB 3061 ATGCTGACATGCTATTTTGTAGGAGAAATATGCTGTAATGATATTTTGAGTTAAA 3120
QY 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTCACATTTTGTACAAATGCTTATGCTTGTA 3180
DB 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTCACATTTTGTACAAATGCTTATGCTTGTA 3180
QY 3181 CAAGCTTATGCTGCTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3240
DB 3181 CAAGCTTATGCTGCTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3240
QY 3241 TGGTTTAGAAAAAGTTAGTATGTGACGATAAATAAATAAATAAATAAATAAATAAATAAATAA 3300
DB 3241 TGGTTTAGAAAAAGTTAGTATGTGACGATAAATAAATAAATAAATAAATAAATAAATAAATAA 3300
QY 3301 TTCAGCACTCCATAAATTTCTATTAATTAATATGCAACATTTTGTGATTTTAAATAAT 3360
DB 3301 TTCAGCACTCCATAAATTTCTATTAATTAATATGCAACATTTTGTGATTTTAAATAAT 3360
```

```
QY 3361 TCCTACTAAGGAATAAAAACTTTAATATACGATATGATATGCTAATAATTTAAAAAGA 3420
DB 3361 TCCTACTAAGGAATAAAAACTTTAATATACGATATGATATGCTAATAATTTAAAAAGA 3420
QY 3421 CATAAATCGATCTCAATTTAGTATTTTAAGATATCTATACTACTATAGGATACAAATCCTACA 3480
DB 3421 CATAAATCGATCTCAATTTAGTATTTTAAGATATCTATACTACTATAGGATACAAATCCTACA 3480
QY 3481 GTTCTCAGATTTACACCTTTTTTTTGTTCATTTGGCTTGTGATCTACACATTTTCCAATCTCTT 3540
DB 3481 GTTCTCAGATTTACACCTTTTTTTTGTTCATTTGGCTTGTGATCTACACATTTTCCAATCTCTT 3540
QY 3541 GCAGCCTCCAGCTCTGGCTTTGCTTACCTGCTGCTTCCCAATGATCTTAATGAAAG 3600
DB 3541 GCAGCCTCCAGCTCTGGCTTTGCTTACCTGCTGCTTCCCAATGATCTTAATGAAAG 3600
QY 3601 TGCAAAAAGAAAAAAGCTTACCAATTAATAAAAAAAGAAAAA 3641
DB 3601 TGCAAAAAGAAAAAAGCTTACCAATTAATAAAAAAAGAAAAA 3641

RESULT 2
AB044088 3641 bp mRNA linear PRI 27-JAN-2001
DEFINITION Homo sapiens mRNA for bHLH protein DEC2, complete, cds.
ACCESSION AB044088
VERSION AB044088.1 GI:12593668
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Fujimoto,K., Shen,M., Noshiro,M., Matsubara,K., Shingu,S.,
Honda,K., Yoshida,E., Suardita,K., Matsuda,Y. and Kato,Y.
TITLE Molecular cloning and characterization of DEC2, a new member of
basic helix-loop-helix proteins
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 164-171 (2001)
MEDLINE 21092582
PUBMED 11162494
REFERENCE 2 (bases 1 to 3641)
AUTHORS Fujimoto,K.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2000) Katsumi Fujimoto, Hiroshima university,
department of biochemistry, school of dentistry; 1-2-3, Kasumi,
Minami-ku, Hiroshima 734-8553, Japan
(E-mail:k.fujimoto@hiroshima-u.ac.jp, Tel:818222575629,
Fax:818222575629)

FEATURES
source
1. 3641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3641
/gene="DEC2"
135..1583
/gene="DEC2"
/codon_start=1
/product="bHLH protein DEC2"
/protein_id="BA521502.1"
/db_xref="GI:1258369"
/translation="MDEGPHLQEROLLEHDFIGLDYSLYNCKPKRSMKRDYDQDT
YKLPRLIIEKRRDRINECTAQKDLLPEHLKLTTLGKLEKAVLVLETLKHLKALFAL
TEQHQKILALQNGRSLSKSPQSDLDAPHSFGQTCAKEVLQYLSRFEWTPPEPCV
LNLHLHVAATQPLPTQLTQOVLPSKGTGAPSAAGSAAPCLERAGOKLEPLAYCV
PVIORTQPAELAENDTDTSGYGGAEARPDREKKGAGASRVITKQPPGEDSPA
PKMKLDSGGSGGGGGGGAAMAALLGPDPAAALLRPDAALLLSLIVAFGGGGG
APPQPAANAAPFLCFPLSPSAANAAYVQPLDKSGLEKLYLPAAAAAPFLLPYRGI
PAPAAAAAATAAATAAAPPCLSSVLSPPPEKAGAAATLLPHEVAPLGAHPHQPH
GRTHLPFAGPREFGNPESQAQEDPSQPGKEAP"

Query Match 100.0%; Score 3641; DB 9; Length 3641;
ORIGIN
```


QY 2161 TCAGCTTCTTAAAGGCATATATTTTCAAGAGAGTGAGATGCGAGTTCTCAGCTTGA 2220
Db 2161 TCAGCTTCTTAAAGGCATATATTTTCAAGAGAGTGAGATGCGAGTTCTCAGCTTGA 2220
QY 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACAC 2280
Db 2221 ACCTATTCTGAAGTGGTTTAAATGGTATCTCTTAGTAACTTGCACCTCGTTAAAGAAACAC 2280
QY 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAGAGAGATGATGAGAGGCCAGATCA 2340
Db 2281 GGAGCTGGGCCATCGTCAGAACTAAGTCAGGAGAGAGATGATGAGAGGCCAGATCA 2340
QY 2341 TTCTTAGTACATTTGCTTAACACTTTTATGAGAAATGACATGAATTAATGACATCATCT 2400
Db 2341 TTCTTAGTACATTTGCTTAACACTTTTATGAGAAATGACATGAATTAATGACATCATCT 2400
QY 2401 TAAATTTCTTAAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTATGCA 2460
Db 2401 TAAATTTCTTAAAGTCCATATATAGATAGATATCTATCTGTACAGATTTCTATTATGCA 2460
QY 2461 TAGATAGGTATCTATACATACATCTCAAGTGCACTTATCCCACTCTCATTAATCCAT 2520
Db 2461 TAGATAGGTATCTATACATACATCTCAAGTGCACTTATCCCACTCTCATTAATCCAT 2520
QY 2521 CATGTTCTTAAATTTTGTAACTTCTACTGTAAAGAAAGTGCACTGAACTTCAAAACAAA 2580
Db 2521 CATGTTCTTAAATTTTGTAACTTCTACTGTAAAGAAAGTGCACTGAACTTCAAAACAAA 2580
QY 2581 ACAAAAACAACAACAACAACAACAAGTCCAACTGATATATCTCTATATCTCTGTAA 2640
Db 2581 ACAAAAACAACAACAACAACAACAAGTCCAACTGATATATCTCTATATCTCTGTAA 2640
QY 2641 TTCAAAGTGAACGAAGCATTTAACTGGCCAGTTTGTATGCAATGCTGTAAAGATAT 2700
Db 2641 TTCAAAGTGAACGAAGCATTTAACTGGCCAGTTTGTATGCAATGCTGTAAAGATAT 2700
QY 2701 AGAATGAAGTCTGTGAGGCTTCTCTATCTCAAGTCTATGATTTCTGGAGACAAAC 2760
Db 2701 AGAATGAAGTCTGTGAGGCTTCTCTATCTCAAGTCTATGATTTCTGGAGACAAAC 2760
QY 2761 CAGATACCAAGATTAACAAAGAGAGCTTTTAAAGGCTTAAACCAAGACCTTGCT 2820
Db 2761 CAGATACCAAGATTAACAAAGAGAGCTTTTAAAGGCTTAAACCAAGACCTTGCT 2820
QY 2821 AGATATTTTGTGTTTCCAAAGTACGCTGTGAGAAATCTCACTGGATGTTATGTA 2880
Db 2821 AGATATTTTGTGTTTCCAAAGTACGCTGTGAGAAATCTCACTGGATGTTATGTA 2880
QY 2881 AGGGGTGACACACAGCTGACTATGATGAGGAGAAATATCTGGCTTTTCTGCTAGT 2940
Db 2881 AGGGGTGACACACAGCTGACTATGATGAGGAGAAATATCTGGCTTTTCTGCTAGT 2940
QY 2941 TTGGTGCATTTGCTGCTGTTGCTTACTGTTTGGCTCAAGCGCTGTGTTAAACAAAGT 3000
Db 2941 TTGGTGCATTTGCTGCTGTTGCTTACTGTTTGGCTCAAGCGCTGTGTTAAACAAAGT 3000
QY 3001 TAAACTCTTAGCTACAGGTGGCTTTATGATGATAGTGTGTAATATCAATCAATTAATG 3060
Db 3001 TAAACTCTTAGCTACAGGTGGCTTTATGATGATAGTGTGTAATATCAATCAATTAATG 3060
QY 3061 ATGCTGTGATGCTATTTTGTAGGAGAGAAATATGCTATGATATTTTGAGTTAA 3120
Db 3061 ATGCTGTGATGCTATTTTGTAGGAGAGAAATATGCTATGATATTTTGAGTTAA 3120
QY 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTGCACTTTTGTACAAATGCTTATGCTGTA 3180
Db 3121 TATCTTTTGGGAGGATTTGCTGAAAGTTGCACTTTTGTACAAATGCTTATGCTGTA 3180
QY 3181 CAGGCTTATGCTCTTAAATATTTTAAAGAAATTAATATCTGCTGTGAGAAACGAGC 3240
Db 3181 CAGGCTTATGCTCTTAAATATTTTAAAGAAATTAATATCTGCTGTGAGAAACGAGC 3240

QY 3241 TGGTTTAGAAAGTTTGTATGTGACGATAACTAGAAATACCTTTATATCTTAGTATT 3300
Db 3241 TGGTTTAGAAAGTTTGTATGTGACGATAACTAGAAATACCTTTATATCTTAGTATT 3300
QY 3301 TTCAGCCTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTAAAAAT 3360
Db 3301 TTCAGCCTCCATAAAATTTCTATTACCTAAATATTGCCACACTATTTTGTGATTAAAAAT 3360
QY 3361 TCTTACTAAGGAATAAAACCTTTTAATATAGATATGATATGCTATATTAATTAAGAA 3420
Db 3361 TCTTACTAAGGAATAAAACCTTTTAATATAGATATGATATGCTATATTAATTAAGAA 3420
QY 3421 CATATGATGCTCAATTTAGTATTTAAGATATCTATACTATAGGATACAAATCACTACA 3480
Db 3421 CATATGATGCTCAATTTAGTATTTAAGATATCTATACTATAGGATACAAATCACTACA 3480
QY 3481 GTTCTCAGATTTACACCTTTTGTGTCATTTGGCTTGATGTCACATTTCCATCTCTT 3540
Db 3481 GTTCTCAGATTTACACCTTTTGTGTCATTTGGCTTGATGTCACATTTCCATCTCTT 3540
QY 3541 GCAAGCTCCAGGCTCTGGCTTTGCTACCTGCTGCTCCCAATGATCTTAATGAAAG 3600
Db 3541 GCAAGCTCCAGGCTCTGGCTTTGCTACCTGCTGCTCCCAATGATCTTAATGAAAG 3600
QY 3601 TGCAAAAAGAAAACCTACCAATTAATAAAAAA 3641
Db 3601 TGCAAAAAGAAAACCTACCAATTAATAAAAAA 3641

RESULT 3
BC025968
LOCUS
DEFINITION Homo sapiens basic helix-loop-helix domain containing, class B, 3,
mRNA (cdna clone MGC:39365 IMAGE:4650325), complete cds.
ACCESSION BC025968
VERSION BC025968.1 GI:19684063
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3431)
AUTHORS Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scherz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
 Parvaneh Saeedi, JR Santos, Angeliqne Schnerch, Ursula Skalska,
 Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacques
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL plate: 42 Row: f Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: GenomesScan gene
 prediction.

FEATURES
 source
 1. .3431
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:39365 IMAGE:4650325"
 /tissue_type="placenta, choriocarcinoma"
 /clone_lib="NIH_MGC_21"
 /lab_host="DH10B-R"
 /notes="Vector: POTB7"
 1. .3431
 /gene="BHLHB3"
 /note="synonyms: SHARP-1, DEC2, SHARPI"
 /db_xref="LocusID:79365"
 /db_xref="MIM:606200"
 292. .1740
 /gene="BHLHB3"
 /codon_start=1
 /product="basic helix-loop-helix domain containing, class
 B_3"
 /protein_id="AAH25968.1"
 /db_xref="GI:19684064"
 /db_xref="LocusID:79365"
 /db_xref="MIM:606200"
 /translation="MDEGPHLQERLLEHDFIGLDYSSLYMCKPKRSMKRDPTKOT
 YKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLERAVVLELTLELKHKALTAL
 TQOQHKITATLONGERSLKSPISQDLDAFHSGFQCAKEVLQYLSRFESWTPREPRCV
 QLNLHLHVAITQFLTPQLLTQOVLSTGTGAPSAAGSAAAPCLERAGSKLEPLAYCV
 PVIRTQPSAEALAEINDITDSGYGEAEAREDRKGGAGNSRTVTKQEPGDSPA
 PKRMKLDSSGGSGGGPGGGAVALAAALIGPDFAAALLRPDAALLSLVAFGGGG
 APFPDAAAAPFCILPFCFLSPSAAAAYVPELDKSLKLYLPAAAPFLPLYPGI
 PAPAAAAAATAAAAAAFAFCLSSVLSPPPPEKAGAAATLLPHEVAPLGAPHQHPH
 GRTHLPFAGPREPGNPESQAQEDPPQPGKEAP"

ORIGIN
 Query Match 83.6%; Score 3043; DB 9; Length 3431;
 Best Local Similarity 99.9%; P: 0; Mismatches 1; Indels 1; Gaps 1;
 Matches 3213; Conservative

QY 1 CTGCACCTGAAGAGGAGCGAGAGAGACTGGAGACGACAGATCCCGCCCAAGTCTC 60
 DB 158 CTGCACCTGAAGAGGAGCGAGAGAGACTGGAGACGACAGATCCCGCCCAAGTCTC 217
 QY 61 CCAAGCCTACCGTCCCAAGATATTGTACAGAGCCCAAAATCGAAACAGAGAAACG 120
 DB 218 CCAAGCCTACCGTCCCAAGATATTGTACAGAGCCCAAAATCGAAACAGAGAAACG 277
 QY 121 AACAGCAGTTGAACATGGACGAGAGGAAATTCCTATTGCAAGAGACAGAGTACTGGAAC 180
 DB 278 AACAGCAGTTGAACATGGACGAGAGGAAATTCCTATTGCAAGAGACAGAGTACTGGAAC 337

QY 181 ATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTATTAACCCAAAGAGCA 240
 DB 338 ATAGAGATTTTATAGGACTGGACTATTCTCTTTGTATATGTATTAACCCAAAGAGCA 397
 QY 241 TGAAGCAGAGCAGACCAAGGATCTTACAAATACCGCAGATTAATAGAAAGAAA 300
 DB 398 TGAAGCAGAGCAGACCAAGGATCTTACAAATACCGCAGATTAATAGAAAGAAA 457
 QY 301 GAAGAGACCGAATTAATGAATGCAATTCAGTGAAGATTTACTGCTGCAACATCTCA 360
 DB 458 GAAGAGACCGAATTAATGAATGCAATTCAGTGAAGATTTACTGCTGCAACATCTCA 517
 QY 361 AATTGCAACTCTGGACATCTGGAGAAAGCTGTAGCTTGGAAATTAATTTGAAACACT 420
 DB 518 AATTGCAACTCTGGACATCTGGAGAAAGCTGTAGCTTGGAAATTAATTTGAAACACT 577
 QY 421 TAAAGCTTTAACCGGCTTAAACCGCAACAGCATCAGAAAGATAATTGCTTTACAGATG 480
 DB 578 TAAAGCTTTAACCGGCTTAAACCGCAACAGCATCAGAAAGATAATTGCTTTACAGATG 637
 QY 481 GGGAGGATCTCTGAATTCGCCCATTCAGTCCGACTTGGATCGTTCCACTCGGAATTC 540
 DB 638 GGGAGGATCTCTGAATTCGCCCATTCAGTCCGACTTGGATCGTTCCACTCGGAATTC 697
 QY 541 AAACATGGCCCAAGAGTCTTGCATATCTCTCCGGTTTGGAGCTGGACACCCAGGG 600
 DB 698 AAACATGGCCCAAGAGTCTTGCATATCTCTCCGGTTTGGAGCTGGACACCCAGGG 757
 QY 601 AGCGCGGTGTGTCCAGCTGATCAACCACTTGCACCGCTGGCCACCCAGTTCTTSCCA 660
 DB 758 AGCGCGGTGTGTCCAGCTGATCAACCACTTGCACCGCTGGCCACCCAGTTCTTSCCA 817
 QY 661 CCCCGCAGCTGTGACTCAACAGTCTCTGAGCAAGAGCACCGGCTTCTCTCGCGG 720
 DB 818 CCCCGCAGCTGTGACTCAACAGTCTCTGAGCAAGAGCACCGGCTTCTCTCGCGG 877
 QY 721 CCGGGTCCGGCGCGCCCTGCTGGAGCGCGGGGAGAGCTGAGAGCCCTCTCGCT 780
 DB 878 CCGGGTCCGGCGCGCCCTGCTGGAGCGCGGGGAGAGCTGAGAGCCCTCTCGCT 937
 QY 781 ACTCGGTGCCCCGTATCCAGCGGACTCAGCCAGCGCGGAGCTGCGCGCGGAGAGCA 840
 DB 938 ACTCGGTGCCCCGTATCCAGCGGACTCAGCCAGCGCGGAGCTGCGCGCGGAGAGCA 997
 QY 841 CGGACACCGACAGCGGCTACGGCGGAGAGCGCGCGGAGCGCGCGGAGAGAGCA 900
 DB 998 CGGACACCGACAGCGGCTACGGCGGAGAGCGCGCGGAGCGCGCGGAGAGAGCA 1057
 QY 901 AAGCGCGGGGCGAGCGCGCTCACCATCAAGCAGAGAGCTCCCGGGAGGAGCTCGCGG 960
 DB 1058 AAGCGCGGGGCGAGCGCGCTCACCATCAAGCAGAGAGCTCCCGGGAGGAGCTCGCGG 1117
 QY 961 CGCCCAAGAGATGAAGTGGATTCGGCGCGCGCGGAGCGCGCGCGCGCGCGCGG 1020
 DB 1118 CGCCCAAGAGATGAAGTGGATTCGGCGCGCGCGGAGCGCGCGCGCGCGCGCGG 1177
 QY 1021 GCGCGCGCGCGGAGCGCGCTTCTGGGGGCCCAACCTGCGCGCGCGCGCGCGCTGC 1080
 DB 1178 GCGCGGTGGGGGAGCGCGCTTCTGGGGGCCCAACCTGCGCGCGCGCGCGCTGC 1237
 QY 1081 TGAGACCCGACGCGCGCTGCTCAGTCTGTGTGGCTTTCGCGGAGCGGAGGCGCG 1140
 DB 1238 TGAGACCCGACGCGCGCTGCTCAGTCTGTGTGGCTTTCGCGGAGCGGAGGCGCG 1297
 QY 1141 CCTTCCGAGCGCGCGCGCGCTTCTGCTGCTTCTGCTTCTTCTCTCTCTCG 1200
 DB 1298 CCTTCCGAGCGCGCGCGCGCGCTTCTGCTGCTTCTGCTTCTTCTCTCTCG 1357
 QY 1201 CTTCTGAGCTGCGCGCTACGTACGCGCTTCTGAGCAAGAGCGGCTTGGAGAGTATC 1260
 DB 1358 CTTCTGAGCTGCGCGCTACGTACGCGCTTCTGAGCAAGAGCGGCTTGGAGAGTATC 1417
 QY 1261 TGTACCGCGCGGCTGCGCGCTTCTGCTGCTTCTATACCCCGGCTATCCCGCGCGG 1320

Db	1418	TGTAACCGCGGCGGCTGCGCGCCCGCTTCGCGCTGCTATACCGCGGATCCCCGCGCCGG	1477
Qy	1321	CGGAGAGCGGCGAGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCTGCCCTGCGCTGT	1380
Db	1478	CGGAGCGCGGCGAGCGCGCGCGCTGCGCGCGCGCGCGCGCTGCCCTGCGCTGT	1537
Qy	1381	CCTCGGTGTGTGCGCGCGCTCCCGAGAGAGGGGCGCGCGCGCGCGACCCCTCTGCGCG	1440
Db	1538	CCTCGGTGTGTGCGCGCGCTCCCGAGAGAGGGGCGCGCGCGCGCGACCCCTCTGCGCG	1597
Qy	1441	ACGAGGTGGGCGCCCTTTGGGCGCGCGACCCCGACCGCGACCGCGCGCACCCACCTGC	1500
Db	1598	ACGAGGTGGCGCCCTTTGGGCGCGCGACCCCGACCGCGACCGCGCGCACCCACCTGC	1657
Qy	1501	CCTTCGCGGGCGCGCGAGCGCGGGGAACCGCGAGAGCTGTCTCAGAGATGCCCTCGC	1560
Db	1658	CCTTCGCGGGCGCGCGAGCGCGGGGAACCGCGAGAGCTGTCTCAGAGATGCCCTCGC	1717
Qy	1561	AGCCAGGAAGGAGCTCCCTGAATCTTGGCTCCCGAAGAGCGAGAGTTCAAGCAGAGT	1620
Db	1718	AGCCAGGAAGGAGCTCCCTGAATCTTGGCTCCCGAAGAGCGAGAGTTCAAGCAGAGT	1777
Qy	1621	GAGAAGTTAAATACCCTTAAGGAGTTCAAGCAGAGTGAGAAAGTTAAAAATACCTTTAAG	1680
Db	1778	GAGAAGTTAAATACCCTTAAGGAGTTCAAGCAGAGTGAGAAAGTTAAAAATACCTTTAAG	1837
Qy	1681	GTCTTTAAGGAGGAGTGTAATAGATGCACGACGAGCATAAACAGAAACAAACAACAG	1740
Db	1838	GTCTTTAAGGAGGAGTGTAATAGATGCACGACGAGCATAAACAGAAACAAACAACAG	1897
Qy	1741	GTGTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCGCGACCACCCACCCCTCCACA	1800
Db	1898	GTGTATGTGTACATTCGGAGTTCCTGTTTGTCTCATCCGCGACCACCCACCCCTCCACA	1957
Qy	1801	CACTAAACATCGCTTCTCCGCCACACAGCTGTAAAGATCTATGCGAAGACACTGGC	1860
Db	1958	CACTAAACATCGCTTCTCCGCCACACAGCTGTAAAGATCTATGCGAAGACACTGGC	2017
Qy	1861	TCCTTTTTTAAATCCCCCAATAAATTTGCCCTTTAGGCGCATGTGCCATTCTCT	1920
Db	2018	TCCTTTTTTAAATCCCCCAATAAATTTGCCCTTTAGGCGCATGTGCCATTCTCT	2077
Qy	1921	TAAATTTGGAACCTAATTCGAGAGGAAGTGAAGAGGTCTGTCTGFGCTGAGCTAGGT	1980
Db	2078	TAAATTTGGAACCTAATTCGAGAGGAAGTGAAGAGGTCTGTCTGFGCTGAGCTAGGT	2137
Qy	1981	GAACCCCGGCTAGGGAAAGATCTTAAACCTTTGACGCTTTTGGAGTTGACATGGAAC	2040
Db	2138	GAACCCCGGCTAGGGAAAGATCTTAAACCTTTGACGCTTTTGGAGTTGACATGGAAC	2197
Qy	2041	AGCAGTGTGTGTATGAGCTAGTTCCTCAAGCTGCCCTGCCTGTTTTAGGAGCGT	2100
Db	2198	AGCAGTGTGTGTATGAGCTAGTTCCTCAAGCTGCCCTGCCTGTTTTAGGAGCGT	2257
Qy	2101	TCACAACACAGATTGAGGCTCTTTTGAATTCGAATTTACTCTTCAGTATTTTCTAATGT	2160
Db	2258	TCACAACACAGATTGAGGCTCTTTTGAATTCGAATTTACTCTTCAGTATTTTCTAATGT	2316
Qy	2161	TCAGCTTTCTAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTTGCA	2220
Db	2317	TCAGCTTTCTAAAGGCATATATTTTCAAAGAGTGAGGATGCAGTTTCTCACGTTGCA	2376
Qy	2221	ACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTGCACTCGTTTAAAGAACAC	2280
Db	2377	ACCTATTCTGAAGTGTTTAAATGGTATCTCTTAGTAACTTGCACTCGTTTAAAGAACAC	2436
Qy	2281	GGAGCTGGGCATCGCTCAGAACTAAGTCAGGAGGAGGATGGAATGAGAGGCCAGAAATCA	2340
Db	2437	GGAGCTGGGCATCGCTCAGAACTAAGTCAGGAGGAGGATGGAATGAGAGGCCAGAAATCA	2496
Qy	2341	TTCCTAGTACATTTGCTAAACATTTTATGTAGAAATTCACCATGAATTAATGGACTCATCT	2400

Db	2497	TTCTGTAGTACATTTTGCTTAAACACCTTTTATGTGAAAAATTGACCATGAAATTAATATGGACTCACT	2555
Qy	2401	TAATTTCTTTCTAAAGTCCATATATAGATAGATATCTATCTGTACAGATTTTCTATTTATCCCA	2460
Db	2557	TAATTTCTTCTAAGTCCATATATAGATAGATATCTATCTGTACAGATTTTCTATTTATCCCA	2616
Qy	2461	TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCATCTCTCATTTAATCCAT	2520
Db	2617	TAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCCATCTCTCATTTAATCCAT	2676
Qy	2521	CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACTGAACTTTCAAAACAAA	2580
Db	2677	CATGTTCTCTAAATTTTGTAAATCTTACTGTAAAAAAGTGCACTGAACTTTCAAAACAAA	2736
Qy	2581	ACAAAAACAACAACAACAACAACAACAAGTCCAAACTGATATATCTCTGTTAAAAA	2640
Db	2737	ACAAAAACAACAACAACAACAACAACAAGTCCAAACTGATATATCTCTGTTAAAAA	2796
Qy	2641	TTCAAAAGTGAACGAAGACATTTAACTGGCCAGCTTTGATTGCAAAATGCTGTTAAAGATAT	2700
Db	2797	TTCAAAAGTGAACGAAGACATTTAACTGGCCAGCTTTGATTGCAAAATGCTGTTAAAGATAT	2856
Qy	2701	AGAATGAAGTCCTGTGAGGCCTTCCATCTCCTCAAGTCTATGTATTTTCTGGAGACCAAAAC	2760
Db	2857	AGAATGAAGTCCTGTGAGGCCTTCCATCTCCTCAAGTCTATGTATTTTCTGGAGACCAAAAC	2916
Qy	2761	CAGATACCAGATTAATCACAAAGAAAGCTTTTTTAATAAGCTTTAAACCAAGACCTTGCTCT	2820
Db	2917	CAGATACCAGATTAATCACAAAGAAAGCTTTTTTAATAAGCTTTAAACCAAGACCTTGCTCT	2976
Qy	2821	AGATATTTTGTGTTGTCACAAAGTAGCACTGTGAGAAATCTCACTGGATGTTATGTA	2880
Db	2977	AGATATTTTGTGTTGTCACAAAGTAGCACTGTGAGAAATCTCACTGGATGTTATGTA	3036
Qy	2881	AGGGGTGAGACACAACAGTCTGACTATGAGTGAAGAAATATCTGGGCTCTTTTCGTCAGT	2940
Db	3037	AGGGGTGAGACACAACAGTCTGACTATGAGTGAAGAAATATCTGGGCTCTTTTCGTCAGT	3096
Qy	2941	TTGGTGCAATTTGCTGCTGTTGCTACTGTTTGGCTCTCAAAACGCTGTTTAAACCAACGT	3000
Db	3097	TTGGTGCAATTTGCTGCTGTTGCTACTGTTTGGCTCTCAAAACGCTGTTTAAACCAACGT	3156
Qy	3001	TAAACTCTTAGCCTACAAGTGGCTCTTATGTACATAGTTGTTTAATACATCCAAATTAATG	3060
Db	3157	TAAACTCTTAGCCTACAAGTGGCTCTTATGTACATAGTTGTTTAATACATCCAAATTAATG	3216
Qy	3061	ATGCTGACATGCTATTTTGTAGGAGAAATATGCTGAATGATATTTTGGAGTTAAAA	3120
Db	3217	ATGCTGACATGCTATTTTGTAGGAGAAATATGCTGAATGATATTTTGGAGTTAAAA	3276
Qy	3121	TATCTTTTGGGAGGATTTGCTGAAAAGTTGCACTTTTGTACAAATGCTTATGCTTGGTA	3180
Db	3277	TATCTTTTGGGAGGATTTGCTGAAAAGTTGCACTTTTGTACAAATGCTTATGCTTGGTA	3336
Qy	3181	CAAGCTATGCTGCTCTTAATATTTTAAAAAAT	3215
Db	3337	CAAGCTATGCTGCTCTTAATATTTTAAAAAAT	3371
RESULT 4			
AC022509/c			
LOCUS			
DEFINITION			
Homo sapiens 12 BAC RP11-283G6 (Roswell Park Cancer Institute Human			
BAC Library) complete sequence.			
linear PRI 30-NOV-2000			

RESULT 4	AC022509/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
----------	------------	-------	------------	-----------	---------	----------	--------	----------	-----------	---------

Alsbrooks, S. L., Amaratunge, H. C., Are, J. R., Banks, T., Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dederich, S. R., David, R., Davila, M. L., Davis, C., Davy, Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H., Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratochvic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louiseged, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondei, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Minor, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunolu, G., Orsugue, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshtari, N., Sieson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezkyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zortilla, S., Zuchterlapati, R. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 204228)
Worley, K. C.
Direct Submission
Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204228)
Worley, K. C.
Direct Submission
Submitted (30-NOV-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 30, 2000 this sequence version replaced gi:10945701.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Option Info -----
Version: 1.01
Std: bcm-qual-stats.cc, v 1.6 1998/10/27 00:34:03 kdurbin Exp kdurbin \$

Options Specified: xfo
Xgapfile: ./xgap/hadw.final
Outfile: ./xgap/HADW.faf
FA File: ./xgap/HADW.faf
Sequence start: 10
Sequence end: 204237
Optimistic: 0
Use no qual cols: 0

----- Summary Statistics -----
Contig length: 204228
Phrap values in estimate: 203313
Average error rate (BCM-Phrap estimate): 0.000109636
Fraction of Phrap values less than 40 : 0.0179674
Number of consensus changing edits: 37
Number of N's in consensus : 0
Clip start: 11
Clip end: 2905
Num reads missing quality: 174
Num reads invalid quality: 0

----- Low Quality Bases -----

Quality	Position	Surrounding Sequence
4	6263	tttttttttt(t)gagacaggg
7	89744	attaataaac(a)aggcagctg
7	89830	ctggtagttt(g)agaccagctg
7	156088	gacaatcaga(a)atgaacttaa
8	16188	ataactaag(c)taaaagaaaa
8	7146	ccactatcac(a)gaatagcat
8	89632	gtcaggaata(t)gccatctgtt
8	89633	tcaggaataat(g)ccatctgtta
8	89634	caggaataat(g)ccatctgttac
8	89745	attaataaca(a)gggcagctgg
8	89791	cagcattttg(g)agggtgagg
8	89792	agcattttgg(g)gggcagaggt
8	89795	attttggagg(g)ctgaggtggg
8	89829	tctgtagttt(t)gagaccagcc
8	140166	ggggaggggg(g)gagggggagc
8	140167	ggggaggggg(g)aggggggagg
9	6160	aaagaaaaa(t)aaatcttggg
9	7030	ctggagagtg(a)agggaagca
9	7097	gagggagagtc(a)agggaagag

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

9	7099	ggagcgcga (a) gggagaagt	Db	44627	ACGACACCGACAGCGCTACGGCGGCGAAGCCGCGGAGCCCGCGGAGAGGC	44568
9	7106	caagggaga (a) gggacacac	Qy	900	AAAGCGCGGGCGAGCGCGCTCAATCAAGAGAGAGCTCCCGGGAGAGACTCGCCG	959
9	7128	gatctcaga (c) aactccccc	Db	44567	AAAGCGCGGGCGAGCGCGCTCAATCAAGAGAGAGCTCCCGGGAGAGACTCGCCG	44508
9	7130	tgacatgaca (c) ctccccac	Qy	960	CGCGCCCAAGAGAGTGAAGCTGGATTCCCGGGCGGCGAGCGCGCGCGCGGGGC	1019
9	7135	gacaactccc (c) cactatcac	Db	44507	CGCGCCCAAGAGAGTGAAGCTGGATTCCCGGGCGGCGAGCGCGCGCGGGGC	44448
9	7136	gacaactccc (c) cactatcac	Qy	1020	GGCGCGCGGGCGGCGAGCGCGCTTCTGGGGCCGACCCCTGCGCGCGGGCGCGCTG	1079
9	7137	caactccccc (a) ctatcacaa	Db	44447	GGCGCGCGGGCGGCGAGCGCGCTTCTGGGGCCGACCCCTGCGCGCGGGCGCGCTG	44388
9	7138	caactccccc (a) ctatcacaa	Qy	1080	CTGAGACCCGAGCGCGCGCTGCTAGCTCGCTGGTGCCTTCGCGGAGCGGAGCGCG	1139
9	7139	aactccccc (c) tatcacaga	Db	44387	CTGAGACCCGAGCGCGCGCTGCTAGCTCGCTGGTGCCTTCGCGGAGCGGAGCGCG	44328
9	7144	ccsactatc (a) caagaatagc	Qy	1140	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCTCTCG	1199
9	7148	actatcacaa (g) aatagcatgg	Db	44327	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCTCTCG	44268
9	7149	ctatcacaa (a) atagcatggg	Qy	1200	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCTCTCG	1259
9	7150	tatcacaga (a) tagcatgggg	Db	44267	CCCTTCCCGCAGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCTCTCG	44208
9	89628	cccgatcagg (a) aatggccatc	Qy	1260	CTGTACCCCGCGCGCGCTGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1319
9	89628	atggtggctt (a) cagctgtaat	Db	44207	CTGTACCCCGCGCGCGCTGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	44148
9	89769	tggtggctt (c) agctgtaat	Qy	1320	CGCGCAGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1379
9	89793	scatttggg (a) ggcctgaggtg	Db	44147	CGCGCAGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	44088
9	89828	stctggtagt (c) tggagccagc	Qy	1380	TCCTTCGCTGTTGTCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1439
9	89865	aaccatcat (c) tacaataaat	Db	44087	TCCTTCGCTGTTGTCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	44028
10	6159	gaaagggaaa (a) taactcttgg	Qy	1440	CAGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1499
10	6179	gggcccacaa (a) tcaactaagt	Db	44027	CAGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	43968
10	6180	ggccccaaaa (t) cactaaagcta	Qy	1500	CCCTTCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1559
10	6209	aagtgaagct (g) ggaactgctt	Db	43967	CCCTTCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	43908
10	6222	aactgcttgg (g) gcaactgctt	Qy	1560	CAGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1619
10	6234	caaatgctt (c) ctatcttttt	Db	43907	CAGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	43848
10	6542	agccacgctg (c) cctaccgctt	Qy	1620	CGCGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	1679
10	7029	ggtggaagct (g) aaggggaagc	Db	43847	CGCGCAGTGGCGCGCGCTTGGGGCGCGCGCGCGCGCGCGCGCTTCTGCGCTGCCCTTCTCTCTCTCG	43788
10	7057	gtctttacaa (g) gtgagcaggag	Qy	1680	GCTCTTTAAGGGAGGAGTGTATAGATGACACAGAGGAGTAAACAGAAACACAAACA	1739
10	7058	tcctttacaa (g) tggcagggagg	Db	43787	GCTCTTTAAGGGAGGAGTGTATAGATGACACAGAGGAGTAAACAGAAACACAAACA	43728
10	7085	gagagagaga (c) agaggggggg	Qy	1740	GGTGTATGTATGATTCGGAGTTCCTGTTTCTCATCCCGCACCAACCCCTCCAC	1799
10	7086	agagagagac (a) gagggagcgc	Db	43727	GGTGTATGTATGATTCGGAGTTCCTGTTTCTCATCCCGCACCAACCCCTCCAC	43668
10	7096	gagagagac (a) gagggagcgc	Qy	1800	ACACTAACATCCCTTCTTCCCGCACCAACCCCTCCAC	1859
10	7100	gagcgcaaa (g) ggaagaagtgc	Db	43667	ACACTAACATCCCTTCTTCCCGCACCAACCCCTCCAC	43608
10	7101	gagcgcaaa (g) gaggaagtgc	Qy	1860	CTCTTTTTTAAATCCCGCAAAATAAATTTTCCCGCCCTTTTAGGCCATGTTTCCATTTATCTC	1919
10	7102	agcgcaaaag (g) agagagtgc	Db	43607	CTCTTTTTTAAATCCCGCAAAATAAATTTTCCCGCCCTTTTAGGCCATGTTTCCATTTATCTC	43548
10	7103	gcgcaaaag (a) gaagtgc	Qy	1920	TTAAATTTGAACTTAATTCGAGAGAGTGAAGAGGCTGTGTTCTGTGCTGAGCTAGG	1979
10	7104	gcgcaaaag (a) gaagtgc	Db	43547	TTAAATTTGAACTTAATTCGAGAGAGTGAAGAGGCTGTGTTCTGTGCTGAGCTAGG	43488
10	7105	gcgcaaaag (a) gaagtgc	Qy			
10	7131	ctcatgacaa (c) tccccacta	Db			
10	7142	tccccacta (t) cacaagaata	Qy			

Query Match

Best Local Similarity

Matches 3143; Conservative

78.4%; Score 2853; DB 9; Length 204228;

99.9%; Pred. No. 0;

0; Mismatches

1; Indels

2; Gaps

2;

Qy	480	GGGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATGCGCTTCCACTCGGGATT	539
Db	44987	GGGGAGCGATCTCTGAATCGCCCATTCAGTCCGACTTGGATGCGCTTCCACTCGGGATT	44928
Qy	540	CAAAATGCGCCAAAGAGCTTTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGG	599
Db	44927	CAAAATGCGCCAAAGAGCTTTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGG	44868
Qy	600	GAGCGCGGTGTGTCAGCTGATCAACACTTGCACGCGCTGGCCACCCAGTCTTGCC	659
Db	44867	GAGCGCGGTGTGTCAGCTGATCAACACTTGCACGCGCTGGCCACCCAGTCTTGCC	44808
Qy	660	ACCCGCGAGCTTGTACTCAACAGTTCCTCTGAGCAAGGACCCGCGGCTCCCTCGCC	719
Db	44807	ACCCGCGAGCTTGTACTCAACAGTTCCTCTGAGCAAGGACCCGCGGCTCCCTCGCC	44748
Qy	720	GCGGGTCCGCGCGCGCCCTGCTGGAGCGCGGGGCGAGAGCTGGAGCCCTCGCC	779
Db	44747	GCGGGTCCGCGCGCGCCCTGCTGGAGCGCGGGGCGAGAGCTGGAGCCCTCGCC	44688
Qy	780	TACTGCTGCGCGTCAATCAGCGGACTCAGCCAGCGCGAGCTCGCGCGCGAGACGAC	839
Db	44687	TACTGCTGCGCGTCAATCAGCGGACTCAGCCAGCGCGAGCTCGCGCGCGAGACGAC	44628
Qy	840	ACGACACCGCAGCGGCTACGGCGGCGAGCCGAGGCGCGCGCGCGAGAGGC	899

Qy	1980	TGAA	CCCCGGGGTAGGGGAAGATGTTAA	CA	CTTTGACGTC	CTTTGGAGTTGACATGGAA	2039
Db	43487	TGAA	CCCCGGGGTAGGGGAAGATGTTAA	CA	CTTTGACGTC	CTTTGGAGTTGACATGGAA	43428
Qy	2040	CAGC	AGTAGTCTGTTATGTAGAC	TAGTTCT	CA	AAGCTGCCCTGCCCTGTTT	TAGAGGCG 2099
Db	43427	CAGC	AGTAGTCTGTTATGTAGAC	TAGTTCT	CA	AAGCTGCCCTGCCCTGTTT	TAGAGGCG 43368
Qy	2100	TTCC	CAAAAACAGATTCAGGCTCT	TTTTAGAA	TTGAA	TTTACTCTTC	CAGTATTTCTTAATG 2159
Db	43367	TTCC	CAAAAACAGATTCAGGCTC	-TTTTAGAA	TTGAA	TTTACTCTTC	CAGTATTTCTTAATG 43309
Qy	2160	TTCAG	CTTTCTAAAAGCATATAT	TTTTT	CA	AAGAAGTGAGGATG	CAGTTTCTCAGCTTGC 2219
Db	43308	TTCAG	CTTTCTAAAAGCATATAT	TTTTT	CA	AAGAAGTGAGGATG	CAGTTTCTCAGCTTGC 43249
Qy	2220	AACCT	ATTCTGAAGTGGTTTTAAAT	TGTTATCT	CTTTAGTAA	CTTCTTGCA	CTCGTTAAAGAAACA 2279
Db	43248	AACCT	ATTCTGAAGTGGTTTTAAAT	TGTTATCT	CTTTAGTAA	CTTCTTGCA	CTCGTTAAAGAAACA 43189
Qy	2280	CGAG	CTGGGCCATCGTCAGAA	CTAAGTCAGG	AAGGAGAT	GGATGGATG	AAGGCCAGAATC 2339
Db	43188	CGAG	CTGGGCCATCGTCAGAA	CTAAGTCAGG	AAGGAGAT	GGATGGATG	AAGGCCAGAATC 43129
Qy	2340	ATTCC	TAGTACATTTGCTAAC	ACCTTTATTCAG	GAATTCG	ACCATGA	ATTAATGGACTCATC 2399
Db	43128	ATTCC	TAGTACATTTGCTAAC	ACCTTTATTCAG	GAATTCG	ACCATGA	ATTAATGGACTCATC 43069
Qy	2400	TTAAT	TTCTTCTTAAGTCCATAT	TAGATAGAT	CTATCTCTG	TACAGATTTCT	TATTTATTC 2459
Db	43068	TTAAT	TTCTTCTTAAGTCCATAT	TAGATAGAT	CTATCTCTG	TACAGATTTCT	TATTTATTC 43009
Qy	2460	ATAGA	TAGGTATCTATACAT	TACACATCTC	AA	TGTCAT	TATTCACCTCATTAATCCA 2519
Db	43008	ATAGA	TAGGTATCTATACAT	TACACATCTC	AA	TGTCAT	TATTCACCTCATTAATCCA 42949
Qy	2520	TCAT	GTTCCTAAATTTTTTG	TAACTTTTAC	TGTAAAA	AAAAAGTCAC	TGAACTTCAAAAACA 2579
Db	42948	TCAT	GTTCCTAAATTTTTTG	TAACTTTTAC	TGTAAAA	AAAAAGTCAC	TGAACTTCAAAAACA 42889
Qy	2580	AA	CAAAAAACA	CA	CA	CAAAAAACA	CAAGTCCAACTGATATCTCTATATTTCTGTTAAA 2639
Db	42888	AA	CAAAAAACA	CA	CA	CAAAAAACA	CAAGTCCAACTGATATCTCTATATTTCTGTTAAA 42829
Qy	2640	ATTCA	AAAGTGAACGAAAGCAT	TTAACTG	CCCAATTTT	CGCATATG	CTATATCTCTGTTAAA 2699
Db	42828	ATTCA	AAAGTGAACGAAAGCAT	TTAACTG	CCCAATTTT	CGCATATG	CTATATCTCTGTTAAA 42769
Qy	2700	TAGA	TGAAGTCTCTGTGAG	CCCTTCTATCT	CCAAGTCTAT	GTATTTT	TGAGAGCCAAA 2759
Db	42768	TAGA	TGAAGTCTCTGTGAG	CCCTTCTATCT	CCAAGTCTAT	GTATTTT	TGAGAGCCAAA 42709
Qy	2760	CCGAT	ATACAGATAATCA	CA	AA	AAAGCTTTT	TATATAGGCTTAAACCAAGACCTTCTC 2819
Db	42708	CCGAT	ATACAGATAATCA	CA	AA	AAAGCTTTT	TATATAGGCTTAAACCAAGACCTTCTC 42649
Qy	2820	TAGA	TATTTTTTAGTTGTT	TGTTGCCA	AGGTAG	CAC	TGTGAGAAATCTCACITTTGGAATGTTATGT 2879
Db	42648	TAGA	TATTTTTTAGTTGTT	TGTTGCCA	AGGTAG	CAC	TGTGAGAAATCTCACITTTGGAATGTTATGT 42589
Qy	2880	AAGG	GTGAGACACACAG	CTC	GACTATAG	TAGTGAG	AAATATCTCGGGCTTTTTCGTGAG 2939
Db	42588	AAGG	GTGAGACACACAG	CTC	GACTATAG	TAGTGAG	AAATATCTCGGGCTTTTTCGTGAG 42529
Qy	2940	TTTGG	TGCATTTGCTGCTG	TGTTGCTACT	TGTTGCT	CAAA	CGCTGTGTTTAAACAACG 2999
Db	42528	TTTGG	TGCATTTGCTGCTG	TGTTGCTACT	TGTTGCT	CAAA	CGCTGTGTTTAAACAACG 42469
Qy	3000	TTAA	ACTTTAG	CCTACA	AGGTGGCTCTTAT	GTAG	TACATAGTTGTTAATACATCCAAATTAAT 3059
Db	42468	TTAA	ACTTTAG	CCTACA	AGGTGGCTCTTAT	GTAG	TACATAGTTGTTAATACATCCAAATTAAT 42409

QY	3060	GATGTCGACATGCTATTTTTGTAGGGAGAAATAATGCTCAATGATATTTTTGAGTTAAA	3111
Db	42408	GATGTCGACATGCTATTTTTGTAGGGAGAAATAATGCTCAATGATATTTTTGAGTTAAA	42349
QY	3120	ATATCTTTTGGGAGAGATTGCTGAAAAAGTTGCACATTTTGTGTACAATGCTTAATGCTTGCT	3179
Db	42348	ATATCTTTTGGGAGAGATTGCTGAAAAAGTTGCACATTTTGTGTACAATGCTTAATGCTTGCT	42289
QY	3180	ACAAGCTTATGCTGCTTTAAATTTATTTTAAAAAATAAATACTGCTCTGTGAGAACCGAG	3239
Db	42288	ACAAGCTTATGCTGCTTTAAATTTATTTTAAAAAATAAATACTGCTCTGTGAGAACCGAG	42230
QY	3240	CTGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAAAATTACCTTTATATTTCTAGTAT	3299
Db	42229	CTGGTTTAGAAAAAGTTTAGTATGTGACGATAAACTAGAAAAATTACCTTTATATTTCTAGTAT	42170
QY	3300	TTTCAGCACCTCCATAAATTCCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAA	3359
Db	42169	TTTCAGCACCTCCATAAATTCCTATTACCTAAATATTGCCACACTATTTTGTGATTTAAAAA	42110
QY	3360	TTCTTACTAGGAATAAAACCTTTAATATACGNATGATATTGTCTTAATAATTTAAAAAG	3419
Db	42109	TTCTTACTAGGAATAAAACCTTTAATATACGNATGATATTGTCTTAATAATTTAAAAAG	42050
QY	3420	ACATAATGGATGCTCAATAGTTTAAAGTATCTATACTATAGGATACAAATCACTAC	3479
Db	42049	ACATAATGGATGCTCAATAGTTTAAAGTATCTATACTATAGGATACAAATCACTAC	41990
QY	3480	AGTCTCAGATTTACACCTTTTTTTTGTGCATGGCTTGATGTGCACACATTTCCAATCTCT	3539
Db	41989	AGTCTCAGATTTACACCTTTTTTTTGTGCATGGCTTGATGTGCACACATTTCCAATCTCT	41930
QY	3540	TGCAAGCTCCAGGCTCTGCTTTTGTCTACCTGCTCGTTCCCAAATGATCTTTAAATGAAAA	3599
Db	41929	TGCAAGCTCCAGGCTCTGCTTTTGTCTACCTGCTCGTTCCCAAATGATCTTTAAATGAAAA	41870
QY	3600	GTGC AAAAGAAAAACCTACCAATTA 3625	
Db	41869	GTGC AAAAGAAAAACCTACCAATTA 41844	
RESULT 5			
BD176928			
LOCUS	BD176928	3274 bp DNA linear	PAT 16-APR-2003
DEFINITION	Novel bHLH type transcriptional gene, DEC2.		
ACCESSION	BD176928		
VERSION	BD176928.1 GI:30014187		
KEYWORDS	JP 2002300878-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Fujimoto, K., Shin, M. and Kato, Y.		
TITLE	Novel bHLH type transcriptional gene, DEC2		
JOURNAL	Patent: JP 2002300878-A 15-OCT-2002;		
COMMENT	CHUGAI PHARMACEUT CO LTD OS Homo sapiens (human) PN JP 2002300878-A/1 PD 15-OCT-2002 PF 19-AUG-1999 JP 1999233286 PI KATSUMI FUJIMOTO, MEI SHIN, YUKIO KATO PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10, PC C12P21/02, G01N33/15, G01N33/50, G01N33/566//A61K31/711, A61K38/00, PC A61K45/00. PC A61K48/00, A61P43/00, C12P21/08, C12N15/00, C12N5/00, A61K37/02 CC FH key Location/Qualifiers FT CDS (135). .(1271).		
FEATURES	Location/Qualifiers		
source	1..3274		

QY		CGCCGCCGCGCGCGCTTCCCCTGCCTGTGTCCTCGGTGTGTGCGCCCCCTCCCGAAGAGGC	1411
Db	985	CGCGCGCGCGCGCGCTTGCCCTTGCCTGTGTCGTGTTGTGCGCCCCCTCCCGAAGAGGC	1044
QY	1412	GGCGCGCGCGCGCGACCTCCTTCGCGCACAGAGTGGCGCCCCCTTGGGCGCGCGACCC	1471
Db	1045	GGCGCGCGCGCGCGACCTCCTTCGCGCACAGAGTGGCGCCCCCTTGGGCGCGCGACCC	1104
QY	1472	CGAGCACCCGGACCGCGCGCACCACTTCGCGCGCGCGCGCGCGCGCGCGGAACCC	1531
Db	1105	CGAGCACCCGGACCGCGCGCACCACTTCGCGCGCGCGCGCGCGCGCGCGGAACCC	1164
QY	1532	GGAGAGCTCTGCTTCAGGAAGATCCCTCGCAGCGAGGAAGAAGCTCCCTGTAATCCTTGC	1591
Db	1165	GGAGAGCTCTGCTTCAGGAAGATCCCTCGCAGCGAGGAAGAAGCTCCCTGTAATCCTTGC	1224
QY	1592	GTCOCGAAGACGGAGGTTCAAGCAGAGTGAGAAGTTPAAAATACCCTTAAGGAGGTTCAA	1651
Db	1225	GTCOCGAAGACGGAGGTTCAAGCAGAGTGAGAAGTTPAAAATACCCTTAAGGAGGTTCAA	1284
QY	1652	GCAGAGTGAGAAGTTAAAAATACCCTTAAGGAGTCTTTAAGGAGGAAGTGTAATAGATGCAC	1711
Db	1285	GCAGAGTGAGAAGTTAAAAATACCCTTAAGGAGTCTTTAAGGAGGAAGTGTAATAGATGCAC	1344
QY	1712	GACAGGCGATAAACGAAGAACCAACAACAGGTGTTATGTGTACATTCCGAGTTCCTGTGTTTT	1771
Db	1345	GACAGGCGATAAACGAAGAACCAACAACAGGTGTTATGTGTACATTCCGAGTTCCTGTGTTTT	1404
QY	1772	GCTCATCCCCGACCAACCCGACCTCCACACACTAAAGTCCCTTTCCTCCCGCCACAGGCT	1831
Db	1405	GCTCATCCCCGACCAACCCGACCTCCACACACTAAAGTCCCTTTCCTCCCGCCACAGGCT	1464
QY	1832	GTAAGAATCCTATGCGAAAGACACTGGCTCTTTTTTTTAAATCCCCCAATAAAATTTTGC	1891
Db	1465	GTAAGAATCCTATGCGAAAGACACTGGCTCTTTTTTTTAAATCCCCCAATAAAATTTTGC	1524
QY	1892	CCCCCTTTAGGCCATGTTCATTTATCTTTAAATTTGGAACCTTAATTCGAGAGAGTAGTAA	1951
Db	1525	CCCCCTTTAGGCCATGTTCATTTATCTTTAAATTTGGAACCTTAATTCGAGAGAGTAGTAA	1584
QY	1952	GAAGGGTCTGTTCTGTGGCTGAGCTAGGTCAACCCCGGGGTAGGGGAAGATGTTAACAC	2011
Db	1585	GAAGGGTCTGTTCTGTGGCTGAGCTAGGTCAACCCCGGGGTAGGGGAAGATGTTAACAC	1644
QY	2012	CTTTTGAGCTTTGGAGTTGACATGGAACAGCGAGTAGTTGTTANGPAGAGCTAGTTCTC	2071
Db	1645	CTTTTGAGCTTTGGAGTTGACATGGAACAGCGAGTAGTTGTTANGPAGAGCTAGTTCTC	1704
QY	2072	AAAGTCCCTCGCTGTTTTAGGAGGGGTTCCACAAACAGATTGAGGCTCTTTTTTAGAAT	2131
Db	1705	AAAGTCCCTCGCTGTTTTAGGAGGGGTTCCACAAACAGATTGAGGCTCTTTTTTAGAAT	1764
QY	2132	TGAATTTACTCTTCAGTAATTTTCTAATGTTTCAGCTTTCTAAAAAGGCATATATTTTCAA	2191
Db	1765	TGAATTTACTCTTCAGTAATTTTCTAATGTTTCAGCTTTCTAAAAAGGCATATATTTTCAA	1824
QY	2192	GAAGTGAGGATGAGTTTCTCAGTTGCAACCTTATCTGAAGTGGTTTAAATGGTATCTC	2251
Db	1825	GAAGTGAGGATGAGTTTCTCAGTTGCAACCTTATCTGAAGTGGTTTAAATGGTATCTC	1884
QY	2252	TTAGTAACCTTGCACTCGCTTAAAGAAACACCGAGCTGGGCCATCGTCAGAACTAAGTCAGG	2311
Db	1885	TTAGTAACCTTGCACTCGCTTAAAGAAACACCGAGCTGGGCCATCGTCAGAACTAAGTCAGG	1944

LOCUS AK094422 1973 bp mRNA linear PRI 30-JAN-2004
 DEFINITION Homo sapiens cDNA FLJ37103 fls, clone BRACE2019327.
 ACCESSION AK094422
 VERSION AK094422.1 GI:21753480
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Ota.T., Suzuki.Y., Nishikawa.T., Otsuki.T., Sugiyama.T., Irie.R., Wakamatsu.A., Hayashi.K., Sato.H., Nagai.K., Kimura.K., Makita.H., Sekine.M., Obayashi.M., Nishi.T., Shibahara.T., Tanaka.T., Ishii.S., Yamamoto.J., Saito.K., Kawai.Y., Isono.Y., Nakamura.Y., Nagahara.K., Murakami.K., Yasuda.T., Iwayanagi.T., Wagatsuma.M., Shiratori.A., Sudo.H., Hosoiri.T., Kaku.Y., Kodaira.H., Kondo.H., Sugawara.M., Takahashi.M., Kanda.K., Yokoi.T., Furuya.T., Kikkawa.E., Omura.Y., Abe.K., Kamihara.K., Katsuta.N., Sato.K., Tanikawa.M., Yamazaki.M., Ninomiya.K., Ishibashi.T., Yamashita.H., Murakawa.K., Fujimori.K., Tanai.H., Kimata.M., Watanabe.M., Hirakawa.S., Chiba.Y., Ishida.S., Ono.Y., Takiguchi.S., Watanabe.S., Hosoda.M., Horuta.T., Kusano.J., Kanehori.K., Takahashi-Fujii.A., Hara.H., Tanase.T., Nomura.Y., Togiya.S., Komai.F., Hara.R., Takeuchi.K., Arita.M., Imose.N., Musashino.K., Yuuki.H., Oshima.A., Sasaki.N., Aotsuka.S., Yoshikawa.Y., Matsunawa.H., Ichihara.T., Shibata.N., Sano.S., Moriya.S., Momiyama.H., Satoh.N., Takami.S., Terashima.Y., Suzuki.O., Nakagawa.S., Senoh.A., Mizoguchi.H., Goto.Y., Shimizu.F., Wakebe.H., Hishigaki.H., Watanabe.T., Sugiyama.A., Takemoto.M., Kawakami.B., Yamazaki.M., Watanabe.K., Kumagai.A., Itakura.S., Fukuzumi.Y., Fujimori.Y., Komiyama.M., Tashiro.H., Tanigami.A., Fujiwara.T., Ono.T., Yamada.K., Fujii.Y., Ozaki.K., Hirao.M., Ohmori.Y., Kawabata.A., Hikiji.T., Kobatake.N., Inagaki.H., Ikema.Y., Okamoto.S., Okitani.R., Kawakami.T., Noguchi.S., Itoh.T., Shigeta.K., Senba.T., Matsumura.K., Nakajima.Y., Mizuno.T., Morinaga.M., Sasaki.M., Togashi.T., Oyama.M., Hata.H., Watanabe.M., Komatsu.T., Mizushima-Sugano.J., Satoh.T., Shirai.Y., Takahashi.Y., Nakagawa.K., Okumura.K., Nagase.T., Nomura.N., Kikuchi.H., Masuho.Y., Yamashita.R., Nakai.K., Yada.T., Nakamura.Y., Ohara.O., Isogai.T. and Sugano.S.
 Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 14702039
 2
 Ninomiya.K., Wagatsuma.M., Kanda.K., Kondo.H., Yokoi.T., Kodaira.H., Furuya.T., Takahashi.M., Kikkawa.E., Omura.Y., Abe.K., Kamihara.K., Katsuta.N., Sato.K., Tanikawa.M., Yamazaki.M., Sugiyama.T., Irie.R., Otsuki.T., Sato.H., Wakamatsu.A., Ishii.S., Yamamoto.J., Isono.Y., Kawai-Hio.Y., Saito.K., Nishikawa.Y., Kimura.K., Yamashita.H., Matsuo.K., Nakamura.Y., Sekine.M., Kikuchi.H., Murakawa.K., Kanehori.K., Takahashi-Fujii.A., Oshima.A., Sugiyama.A., Kawakami.B., Suzuki.Y., Sugano.S., Nagahara.K., Masuho.Y., Nagai.K. and Isogai.T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1973)
 Isogai.T. and Yamamoto.J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..1973
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="BRACE2019327"
 /tissue_type="cerebellum"
 /clone_lib="BRACE2"
 /notes="cloning vector: pME18SFL3"
 ORIGIN
 Query Match 25.3%; Score 921; DB 9; Length 1973;
 Best Local Similarity 99.8%; Pred. NO. 0;
 Matches 1091; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 2533 TTTTGTGTAATCTTACTGCTAAAAAAGTGCACCTGAACTTCAAAAACAAAAACAAACAAAC 2592
 DB 1 TTTTGTGTAATCTTACTGCTAAAAAAGTGCACCTGAACTTCAAAAACAAAAACAAAC 60
 QY 2593 AACACAAAAAACAAGTCCAACTGATATATCTTATCTGTTAAATTTCAAAAGTGA 2652
 DB 61 AACACAAAAAACAAGTCCAACTGATATATCTTATCTGTTAAATTTCAAAAGTGA 120
 QY 2653 CGAAGCATTTAACTGCGCAGTTTTCATTGCAATCTCTGTAAGATATAGATGAAGTCC 2712
 DB 121 CGAAGCATTTAACTGCGCAGTTTTCATTGCAATCTCTGTAAGATATAGATGAAGTCC 180
 QY 2713 TGTGAGGCTTCTCTATCTCCAGTCTATGATTTTCTGAGACCAACCCAGATACCAGAT 2772
 DB 181 TGTGAGGCTTCTCTATCTCCAGTCTATGATTTTCTGAGACCAACCCAGATACCAGAT 240
 QY 2773 AATCACAAGGAGGCTTTTAAATAGGCTTAAACCAAGACCTTGCTAGATATTTTAG 2832
 DB 241 AATCACAAGGAGGCTTTTAAATAGGCTTAAACCAAGACCTTGCTAGATATTTTAG 300
 QY 2833 TTTGTTGCCAAGTAGCAGTCTGAGAAATCTCACTGGAATGTTATGTAAGGGGTGAGACA 2892
 DB 301 TTTGTTGCCAAGTAGCAGTCTGAGAAATCTCACTGGAATGTTATGTAAGGGGTGAGACA 360
 QY 2893 CAACAGTCTGACCTAGTGAGGAGAAATATCTGGGCTTTTTCGTGCTAGTTGCTGATTTG 2952
 DB 361 CAACAGTCTGACCTAGTGAGGAGAAATATCTGGGCTTTTTCGTGCTAGTTGCTGATTTG 420
 QY 2953 CTGCTGCTGTTGCTACTGTTTGGCTCAACAGCTGCTGTTTAAACCAAGCTTAAACTCTTAC 3012
 DB 421 CTGCTGCTGTTGCTACTGTTTGGCTCAACAGCTGCTGTTTAAACCAAGCTTAAACTCTTAC 480
 QY 3013 CTACAGGCTGGCTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 3072
 DB 481 CTACAGGCTGGCTCTTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 540
 QY 3073 CTATTTTGTAGGAGAAATATGCTAATGATATTTTGGTAAATATCTTTTGGGG 3132
 DB 541 CTATTTTGTAGGAGAAATATGCTAATGATATTTTGGTAAATATCTTTTGGGG 600
 QY 3133 AGGATTTGCTGAAAAGTGCACCTTTTGTACAAGTCTTATGTTGTTGTAACAGCTTATGCT 3192
 DB 601 AGGATTTGCTGAAAAGTGCACCTTTTGTACAAGTCTTATGTTGTTGTAACAGCTTATGCT 660
 QY 3193 GTCTTAAATTTTAAAAAATTAATCTGCTGAGAACCCAGCTGGTTTCAAAA 3252
 DB 661 GTCTTAAATTTTAAAAAATTAATCTGCTGAGAACCCAGCTGGTTTCAAAA 719
 QY 3253 GTTTAGTATGTCAGGATAAAGTAACTGCTTATATTTTATTTTATTTTATTTTATTTTATTTT 3312
 DB 720 GTTTAGTATGTCAGGATAAAGTAACTGCTTATATTTTATTTTATTTTATTTTATTTTATTTT 779
 QY 3313 TAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3372
 DB 780 TAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 839
 QY 3373 ATAAAACTTTAATATACGATATGATATGCTTAATAATTTAAAAAAGACATAATGATGTC 3432
 DB 840 ATAAAACTTTAATATACGATATGATATGCTTAATAATTTAAAAAAGACATAATGATGTC 899
 QY 3433 TCAATTTAGTTTAAAGATATCTTAACTATAGGATACAAATCACTAGTTCTCAGATTTT 3492

FEATURES
 source


```
Db 601 CCCGTCATCCAGCGACTCAGCCAGCGCCGAGCTCGCCGCCGAGAACGACACGACACC 660
QY 849 GACGCGGCTACGGCGGCGAAGCCGAGCGCCGCGCCGCGGAGAGGAAAGGCGCG 908
Db 661 GACGCGGCTACGGCGGCGAAGCCGAGCGCCGCGCCGCGGAGAGGAAAGGCGCG 720
QY 909 GGGGCGAGCCGC 920
Db 721 GGGGCGAGCCGC 732

RESULT 11
HSA270695 593 bp mRNA linear PRI 30-MAY-2001
LOCUS Homo sapiens partial unknown mRNA from drug-resistant melanoma
DEFINITION cells, 3'UTR, clone DMS-11.
ACCESSION AJ270695
VERSION AJ270695.1 GI:8017389
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Grotke,C., Mantwill,K., Dietel,M., Schadendorf,D. and Lage,H.
TITLE Identification of differentially expressed genes in human melanoma
JOURNAL Int. J. Cancer 88 (4), 535-546 (2000)
MEDLINE 20514193
PUBMED 11058868
REFERENCE 2 (bases 1 to 593)
AUTHORS Lage,H.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1999) Lage H., Institute of Pathology, Humboldt
University Berlin, Germany
D-10117 Berlin, GERMANY
FEATURES
source
1..593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DSM-11"
/cell_line="MeWo"
/cell_type="melanoma"
<1..593
3'UTR
polya_site
ORIGIN
Query Match 12.6%; Score 460; DB 9; Length 593;
Best Local Similarity 99.8%; Pred. No. 5.7e-237;
Matches 580; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1978 GGTGAACCCCGGGGTAGGGGAAAGATGTTAACACCTTCAGCTCTTGAGTTGACATGG 2037
Db 7 GGTGAACCCCGGGGTAGGGGAAAGATGTTAACACCTTCAGCTCTTGAGTTGACATGG 66
QY 2038 AACAGCAGGTAGTGTATTGTAGAGCTAGTCTTCAAAGCTGCCCTGCTTTTAGGAGG 2097
Db 67 AACAGCAGGTAGTGTATTGTAGAGCTAGTCTTCAAAGCTGCCCTGCTTTTAGGAGG 126
QY 2098 CGTTCACAAACAGATGAGGCTCTTTTAGAATTGAATTACTCTTCAGTATTTCTTAA 2157
Db 127 CGTTCACAAACAGATGAGGCTC-TTTTAGAATTGAATTACTCTTCAGTATTTCTTAA 185
QY 2158 TGTTCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGATGCGATTTCTCAGCTT 2217
Db 186 TGTTCAGCTTCTTAAAGGCATATATTTTCAAAGAGTGAGATGCGATTTCTCAGCTT 245
QY 2218 GCAACCTATTCTGAAGTGGTTTAAATGGTATCTCTTTAGTAACTTCGACCTCGTTAAAGAA 2277
Db 246 GCAACCTATTCTGAAGTGGTTTAAATGGTATCTCTTTAGTAACTTCGACCTCGTTAAAGAA 305
```

```
QY 2278 CACGGAGCTGGGCCATCGTCAGAACTAAGTCAGGAGGAGATGATGAGAGGCCAGAA 2337
Db 306 CACGGAGCTGGGCCATCGTCAGAACTAAGTCAGGAGGAGATGATGAGAGGCCAGAA 365
QY 2338 TCATTCTAGTACATTTGCTTAACACACTTTATTGAGAAATTGACCATGAATTAATGGAATCA 2397
Db 366 TCATTCTAGTACATTTGCTTAACACACTTTATTGAGAAATTGACCATGAATTAATGGAATCA 425
QY 2398 TCTTAATTTCTTAAAGTCCATATATAGATATCTATCTGTACAGATTTCTATTAT 2457
Db 426 TCTTAATTTCTTAAAGTCCATATATAGATATCTATCTGTACAGATTTCTATTAT 485
QY 2458 CCATAGATAGGTATCATATACATACATCTCAAGTGCATCTATTCCTCATCTCAATATC 2517
Db 486 CCATAGATAGGTATCATATACATACATCTCAAGTGCATCTATTCCTCATCTCAATATC 545
QY 2518 CATCATGTTCTCTAAATTTTGTAAATCTTACTGTGTAATAAAAAA 2558
Db 546 CATCATGTTCTCTAAATTTTGTAAATCTTACTGTGTAATAAAAAA 586

RESULT 12
BV208394 797 bp DNA linear STS 01-JUL-2004
LOCUS BHLH3 2085 Rhesus macaque genomic DNA Macaca mulatta STS genomic
DEFINITION clone MWA2085, sequence tagged site.
ACCESSION BV208394
VERSION BV208394.1 GI:49533077
KEYWORDS STS.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
REFERENCE 1 (bases 1 to 797)
AUTHORS Spindel,E.R., Pauley,M., Jia,Y., Boyle,N., Jiang,S., Gravett,C.,
Lupo,S.L., Ali,H., Ojeda,S.R. and Norgren,R.B.
TITLE Targeted amplification of the 3' end of rhesus macaque orthologs of
human genes
JOURNAL Unpublished (2004)
COMMENT
Contact: Spindel ER
Division of Neuroscience
Oregon National Primate Research Center
505 NW 185th Avenue, Beaverton, OR 97006, USA
Tel: 403-690-5388
Fax: 503-690-5384
Email: spindele@ohsu.edu
Primer A: atgaattaatggactcatctt
Primer B: actaaactttctaaaccagc
STS size: 797
PCR Profile:
Hot Start: 95 degrees C for 2.00 min
Denaturation: 95 degrees C for 0.50 min
Annealing: 51 degrees C for 0.50 min
Polymerization: 72 degrees C for 1.00 min
PCR Cycles: 35
Extension: 72 degrees C for 7.0 min
Thermal Cycler: MJ Instruments PTC100
Protocol:
Template: 200 ng
Primer: each 1uM
dNTP's: each 200 uM
Tag Polymerase: 0.05 units/ul (Fast Start High
Fidelity, Roche)
Total Vol: 50 ul
Buffer: MgCl2: 1.8 mM
Fast Start polymerase reaction buffer (Roche)
```

Bases 156-797 are 97% homologous (Blast) to bases 2596-3238 of NM_030762.1. Primers were chosen to amplify genomic DNA in the 3'

ORIGIN

Query Match	10.6%;	Score 386;	DB 6;	Length 447;
Best Local Similarity	100.0%;	Pred. No. 7.4e-197;		
Matches 386;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2445	GATTTCTATTATTCCATAGATAGGTATCTATACATACACATCTCAAGTGCATCTATTCCC	2504	
DB	62	GATTTCTATTATTCCATAGTAGTATCTATACACATCTCAAGTGCATCTATTCCC	121	
QY	2505	ACTCTCATTAATCCATCATGTTCCTAAATTTTGTATCTTACTGTAAAAAAGTGCAC	2564	
DB	122	ACTCTCATTAATCCATCATGTTCCTAAATTTTGTATCTTACTGTAAAAAAGTGCAC	181	
QY	2565	TGAACCTCAAAAACAAAACAAAACAAACACACAAAAAACCAAGTCCAACTCATATATC	2624	
DB	182	TGAACCTCAAAAACAAAACAAAACAAACACACAAAAAACCAAGTCCAACTCATATATC	241	
QY	2625	CTATATTCTGTCTAAAAATTCAAAAGTGAACGAAAGCATTTTAACTGGCCAGTTTTGATTGCA	2684	
DB	242	CTATATTCTGTCTAAAAATTCAAAAGTGAACGAAAGCATTTTAACTGGCCAGTTTTGATTGCA	301	
QY	2685	AATCGCTGAAGATATAGAATGAAGTCCTGTGAGGCCCTTCCTATCTCCAAGTCTATGTAT	2744	
DB	302	AATCGCTGAAGATATAGAATGAAGTCCTGTGAGGCCCTTCCTATCTCCAAGTCTATGTAT	361	
QY	2745	TTTCTGAGACCAAAACCAAGATACCAGATATCAAGAAGAAAGCTTTTTTAAATAAAGGCTTA	2804	
DB	362	TTTCTGAGACCAAAACCAAGATACCAGATATCAAGAAGAAAGCTTTTTTAAATAAAGGCTTA	421	
QY	2805	AAACCAAGACCTTGCTAGATATTTTT	2830	
DB	422	AAACCAAGACCTTGCTAGATATTTTT	447	

RESULT 15

CQ685084
 LOCUS CQ685084 471 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 30010 from Patent WO02070737.
 ACCESSION CQ685084
 VERSION CQ685084.1 GI:42211631
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
 TITLE Compositions and methods relating to osteoarthritis
 JOURNAL Patent: WO 02070737-A 30010 12-SEP-2002;
 Chondrogene Inc. (CA)
 FEATURES
 source
 1..471
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

	Query Match	10.0%;	Score 363;	DB 6;	Length 471;
	Best Local Similarity	100.0%;	Pred. No. 2.2e-184;		
	Matches 363;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2629	ATTCTGTTAAAAATCAAAGCTGAACGAAAGCATTTAACTGGCCAGATTTTCATTGCAAAATG	2658		
DB	1	ATTCTGTTAAAAATCAAAGCTGAACGAAAGCATTTAACTGGCCAGATTTTCATTGCAAAATG	60		
QY	2689	CTGTAAAGATATAGAAATGAAGTCTGTGAGGCCTTCTATCTCCAAAGTCTATGTATTTTC	2748		
DB	61	CTGTAAAGATATAGAAATGAAGTCTGTGAGGCCTTCTATCTCCAAAGTCTATGTATTTTC	120		
QY	2749	TGAGAGCCAAACAGAGATACAGAGATAATCACAAAGAAGCTTTTTTAATAGGGTTTAAACC	2808		
DB	121	TGAGAGCCAAACAGAGATACAGAGATAATCACAAAGAAGCTTTTTTAATAGGGTTTAAACC	180		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 17, 2004, 04:12:41 ; Search time 4563.79 Seconds
(without alignments)
12064.617 Million cell updates/sec

Title: US-10-078-650-11
Perfect score: 1511
Sequence: 1 catgacgaggaattcttc.....gtgagaagttaataacct 1511

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	32.0	523	1	AI672308
2	434	28.7	715	4	BM542244
3	431	28.5	522	1	AI700204
4	430	28.5	557	1	AI819798
5	387	25.6	526	1	AI459114
6	380	25.1	515	2	BE857082
7	380	25.1	557	2	AW073334
8	322	21.3	434	1	AA996006
9	310	20.5	310	9	CG464925
10	301	19.9	301	1	AI022846
11	299	19.8	299	1	AI028096
12	288	19.1	451	1	AI346780
13	283	18.7	522	2	BF432960
14	276	18.3	1109	3	CR615693
15	262	17.3	349	1	AA983880
16	256	16.9	281	2	AW136718
17	252	16.7	388	2	AW977184
18	237	15.7	292	1	AI094420
19	230	15.2	420	4	BG770004
20	223	14.8	350	1	AI357854
21	212	14.0	356	1	AA828321
22	212	14.0	353	1	AA013045
23	212	14.0	392	1	AA279980
24	205	13.6	464	1	AV727296

25	204	13.5	1056	7	CF552953
26	203	13.4	254	1	AI242295
27	188	12.4	369	1	AA053949
28	182	12.0	390	1	AI538997
29	176	11.6	488	7	R93946
30	176	11.6	892	5	BQ955258
31	171	11.3	791	4	BG741953
32	170	11.3	901	1	AI542834
33	169	11.2	169	1	AI356664
34	159	10.5	224	1	AA013207
35	146	9.7	439	1	AA814860
36	137	9.1	235	1	AA730023
37	133	8.8	188	1	AA807164
38	128	8.5	374	1	AI826340
39	128	8.5	391	5	BU580066
40	128	8.5	431	2	AW449776
41	128	8.5	466	5	BU734221
42	128	8.5	484	4	BG482317
43	128	8.5	521	5	BU622650
44	128	8.5	839	6	CD110008
45	119	7.9	385	5	BU579816

ALIGNMENTS

RESULT 1
AI672308
LOCUS
DEFINITION
ty63f03.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2283773 3', similar to contains TAR1.t2 MSR1 MSR1 repetitive element ;, mRNA sequence.
AI672308 523 bp mRNA linear EST 18-MAY-1999
AI672308.1 GI:4852039
EST.
SOURCE
Homo sapiens (human)

ACCESSION
VERSION
KEYWORDS
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI672308 523 bp mRNA linear EST 18-MAY-1999
ty63f03.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2283773 3', similar to contains TAR1.t2 MSR1 MSR1 repetitive element ;, mRNA sequence.
AI672308
AI672308.1 GI:4852039
EST.
SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40Up from Gibco
High quality sequence stop: 490.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2283773"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: Kidney; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site1: Not 1; Site2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid1 prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M.

FEATURES
source

Seq primer: -40UP from Gibco
High quality sequence stop: 451.
FEATURES
source

1. .522
/organism="Homo sapiens"
/ncl_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2330737"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/notes="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbH19W, testis NHT, and B-cell NCI-CGAP GC81) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 28.5%; Score 431; DB 1; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.5e-211; Mismatches 0; Indels 0; Gaps 0;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 193
DB 431 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 372

QY 194 TGCATTGCTCAGCTGAAGATTACTGCTGACATCTGAAATTCACAACTCTGGACAT 253
DB 371 TGCATTGCTCAGCTGAAGATTACTGCTGACATCTGAAATTCACAACTCTGGACAT 312

QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACTTAAAGCTTTAACCGCCTTA 313
DB 311 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACTTAAAGCTTTAACCGCCTTA 252

QY 314 ACCGAGCAACAGCATCAGAAATTAATGCTTTACAGATGGGAGCGATCTCTGAAATCG 373
DB 251 ACCGAGCAACAGCATCAGAAATTAATGCTTTACAGATGGGAGCGATCTCTGAAATCG 192

QY 374 CCCATTTCAGTCCGACTTGATGCTTCCACTCGGATTTCAACATCGCCCAAGAGTC 433
DB 191 CCCATTTCAGTCCGACTTGATGCTTCCACTCGGATTTCAACATCGCCCAAGAGTC 132

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGTCTCCAGCTG 493
DB 131 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGTCTCCAGCTG 72

QY 494 ATCAACGACTTCAGCCGCTGGCCACCGAGTTCTTCCACCCGCGAGCTGTGACTCAA 553
DB 71 ATCAACGACTTCAGCCGCTGGCCACCGAGTTCTTCCACCCGCGAGCTGTGACTCAA 12

QY 554 CAGGTCCCTCT 564
DB 11 CAGGTCCCTCT 1

RESULT 4

AI819798/c
LOCUS
DEFINITION
wJ45a04.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2405742 3,
similar to TR:035779 035779 ENHANCER-OP-SPLIT AND HAIRY-RELATED
PROTEIN 1. ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AI819798
AI819798.1 GI:5438877
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (Bases 1 to 557)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 340 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 441.

FEATURES
source

1. .557
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2405742"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="NCI CGAP Lu19"
/clone_lib="NCI CGAP Lu19"
/notes="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dR) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

ORIGIN

Query Match 28.5%; Score 430; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 5.1e-211; Mismatches 0; Indels 0; Gaps 0;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 193
DB 431 GATACCTACAAATTACCGCACAGATTAAATAGAAAAGAAAGAGAGACCGAATTAATGAA 372

QY 194 TGCATTGCTCAGCTGAAGATTACTGCTGACATCTGAAATTCACAACTCTGGACAT 253
DB 371 TGCATTGCTCAGCTGAAGATTACTGCTGACATCTGAAATTCACAACTCTGGACAT 312

QY 254 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACTTAAAGCTTTAACCGCCTTA 313
DB 311 CTGGAGAAAGCTGTAGTCTTGGAAATTAATTTGAAACACTTAAAGCTTTAACCGCCTTA 252

QY 314 ACCGAGCAACAGCATCAGAAATTAATGCTTTACAGATGGGAGCGATCTCTGAAATCG 373
DB 251 ACCGAGCAACAGCATCAGAAATTAATGCTTTACAGATGGGAGCGATCTCTGAAATCG 192

QY 374 CCCATTTCAGTCCGACTTGATGCTTCCACTCGGATTTCAAAACATCGCCCAAGAGTC 433
DB 191 CCCATTTCAGTCCGACTTGATGCTTCCACTCGGATTTCAAAACATCGCCCAAGAGTC 132

QY 434 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGTCTCCAGCTG 493
DB 131 TTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGAGCCGCGGTGTCTCCAGCTG 72

QY 494 ATCAACGACTTCAGCCGCTGGCCACCGAGTTCTTCCACCCGCGAGCTGTGACTCAA 553
DB 71 ATCAACGACTTCAGCCGCTGGCCACCGAGTTCTTCCACCCGCGAGCTGTGACTCAA 12

```

QY 554 CAGGTCCCTC 563
Db 11 CAGGTCCCTC 2

RESULT 5
AI459114
LOCUS
DEFINITION
t165c02.x1 Soares NSF F8 9W OT PA P S1 Homo sapiens cDNA clone
IMAGE:2146370 3' similar to contains TAR1.t2 MSR1 repetitive
element ; mRNA sequence.
AI459114
VERSION AI459114.1 GI:4311693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 526)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 653 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 452.
FEATURES
Location/Qualifiers
1..526
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2146370"
/lab_host="DH10B"
/notes="Organ: Soares NSF F8 9W OT PA P S1"
a modified polylinker; Vector: p7T3D-Pac (Pharmacia) with
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSP pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-728407, 739050-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 25.6%; Score 387; DB 1; Length 526;
Best Local Similarity 99.8%; Pred. No. 9.8e-189;
Matches 437; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1074 CTTCGAGCTGCGGCTAGTCAGCCCTTCTTGACAGAGCGGCGCTGGAGAGTATC 1133
Db 1 CTTCGAGCTGCGGCTAGTCAGCCCTTCTTGACAGAGCGGCGCTGGAGAGTATC 60

QY 1134 TGTACCGGGGGGGTGGCGGCGGCTTCCCGTGTATACCCCGCATCCCGCGCCCGG 1193
Db 61 TGTACCGGGGGGGTGGCGGCGGCTTCCCGTGTATACCCCGCATCCCGCGCCCGG 120

QY 1194 CGGACGCGCGGAGCGCGCGCGCGCTGCGCGCGCGCGCGGTTTCCCTGCGCTGT 1253
Db 121 CGGACGCGCGGAGCGCGCGCGCGCTGCGCGCGCGCGCGGTTTCCCTGCGCTGT 180

```

```

QY 1254 CCTCGGTGTTGTCGCCCCCTCCCGAGAGCGCGCGCGCGCGCCCTCTCTGCGC 1313
Db 181 CCTCGGTGTTGTCGCCCCCTCCCGAGAGCGCGCGCGCGCGCCCTCTCTGCGC 240

QY 1314 ACAGAGTGGCGCCCCCTTGGGGCGCGCACCCCCAGACACCGCGCGCGCCACTGC 1373
Db 241 ACAGAGTGGCGCCCCCTTGGGGCGCGCACCCCCAGACACCGCGCGCGCCACTGC 300

QY 1374 CCTTCGCGCGCGCGCGCGCGCGCGGAAACCGGAGAGCTCTGTCTAGGAAGATCCCTGC 1433
Db 301 CCTTCGCGCGCGCGCGCGCGCGGAAACCGGAGAGCTCTGTCTAGGAAGATCCCTGC 360

QY 1434 AGCCAGGAAAGAGAGCTCCCTGATCCTTGGTCCCGAAGACGAGGTTCAAGCAGAGT 1493
Db 361 AGCCAGGAAAGAGAGCTCCCTGATCCTTGGTCCCGAAGACGAGGTTCAAGCAGAGT 420

QY 1494 GAGAAGTTAAATACCT 1511
Db 421 GAGAAGTTAAATACCT 438

RESULT 6
BE857082/c
LOCUS
DEFINITION
7923f11.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3307341 3'
similar to TR:Q35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
PROTEIN 1. ; mRNA sequence.
ACCESSION BE857082
VERSION BE857082.1 GI:10370753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 515)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BrTAP), Tumor Gene Index
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 459.
FEATURES
Location/Qualifiers
1..515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3307341"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/notes="Organ: Brain; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo (dT) primer [5'
TGTACCACTCTCAAGTGGAGCGCGCGCATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and EcoRI sites of the modified p7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

```

```
Query Match      25.1%; Score 380; DB 2; Length 515;
Best Local Similarity 99.8%; Pred. No. 4.2e-185;
Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 134 GATACCTACAAATACCGCACAGATTAAAGAAAAGAAAGAGACCGCAATTAATGAA 193
Db 431 GATACCTACAAATACCGCACAGATTAAAGAAAAGAAAGAGACCGCAATTAATGAA 372
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 253
Db 371 TGCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 312
QY 254 CTGAGAAAGCTGTAGCTTTGGAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCTTA 313
Db 311 CTGAGAAAGCTGTAGCTTTGGAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCTTA 252
QY 314 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGAGGATCTCTGAATCG 373
Db 251 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGAGGATCTCTGAATCG 192
QY 374 CCCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 433
Db 191 CCCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 372
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGGAGCGCGGTGTCTGAGCTG 493
Db 131 TTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGGAGCGCGGTGTCTGAGCTG 72
QY 494 ATCAACCACTTTGACGCGGTGGACCCAGCTTCTGCGCCACCCGAGCTGTGTGACTCA 553
Db 71 ATCAACCACTTTGACGCGGTGGACCCAGCTTCTGCGCCACCCGAGCTGTGTGACTCA 12
QY 554 CAGGTCCCTCT 564
Db 11 CAGGTCCCTCT 1
```

```
RESULT 7
AW073334/c
LOCUS
DEFINITION
  wy97905.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2556536 3'
  similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
  PROTEIN 1. ; mRNA sequence.
ACCESSION
  AW073334
VERSION
  AW073334.1 GI:6028332
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 557)
REFERENCE
  NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute / National Institute of Neurological
  Disorders and Stroke, Brain Tumor Genome Anatomy Project
  (CGAP/BRGAP), Tumor Gene Index
  Unpublished (1998)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
  Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LINL at:
  www-bio.lnl.gov/bbrp/image/image.html
  Seq primer: -40UP from Gibco
  High quality sequence stop: 455.
  Location/Qualifiers
  1..557
  /organism="Homo sapiens"
  /mol_type="mRNA"
```

```
ORIGIN
Query Match      25.1%; Score 380; DB 2; Length 557;
Best Local Similarity 99.8%; Pred. No. 4.2e-185;
Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 134 GATACCTACAAATACCGCACAGATTAAAGAAAAGAAAGAGACCGCAATTAATGAA 193
Db 431 GATACCTACAAATACCGCACAGATTAAAGAAAAGAAAGAGACCGCAATTAATGAA 372
QY 194 TGCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 253
Db 371 TGCATTGCTCAGCTGAAAGATTACTGCTGAACTGAAATTTGACAACTCTGGGACAT 312
QY 254 CTGAGAAAGCTGTAGCTTTGGAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCTTA 313
Db 311 CTGAGAAAGCTGTAGCTTTGGAATTAACCTTTGAAACACTTTAAAGCTTTAAACCGCTTA 252
QY 314 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGAGGATCTCTGAATCG 373
Db 251 ACCGAGCAACAGCATCAGAGATTAATGCTTTACAGATGGGAGGATCTCTGAATCG 192
QY 374 CCCATTGCTCAGCTTGGATTCGTTCCATCTCGGGATTTCAAACATGCGCCAAAGAGTC 433
Db 191 CCCATTGCTCAGCTTGGATTCGTTCCATCTCGGGATTTCAAACATGCGCCAAAGAGTC 132
QY 434 TTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGGAGCGCGGTGTCTGAGCTG 493
Db 131 TTGCAATACCTCTCCCGGTTTGAGAGTGGACACCCAGGAGCGCGGTGTCTGAGCTG 72
QY 494 ATCAACCACTTTGACGCGGTGGACCCAGCTTCTGCGCCACCCGAGCTGTGTGACTCA 553
Db 71 ATCAACCACTTTGACGCGGTGGACCCAGCTTCTGCGCCACCCGAGCTGTGTGACTCA 12
QY 554 CAGGTCCCTCT 564
Db 11 CAGGTCCCTCT 1

RESULT 8
AA996006/c
LOCUS
DEFINITION
  os26c04.s1 NCI_CGAP_Kids Homo sapiens cDNA clone IMAGE:1606470 3'
  similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
  PROTEIN 1. ; mRNA sequence.
ACCESSION
  AA996006
VERSION
  AA996006.1 GI:3182495
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 434)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
```



```
RESULT 10
AI022846/c 301 bp mRNA linear EST 28-AUG-1998
LOCUS
DEFINITION
Ow54411.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGE:1650620 3', similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND
HAIRY-RELATED PROTEIN 1. ; mRNA sequence.

ACCESSION
AI022846
VERSION
AI022846.1 GI:3238087
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 301)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 317 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES
Location/Qualifiers
1..301
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1650620"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="PH10B (ampicillin resistant)"
/clone_lib="Soares parathyroid tumor NbHPA"
/notes="Organ: parathyroid gland; Vector: pT7T3D
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCGACCAATTTTTTTTTTTTTTTT
TTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

ORIGIN
Query Match 19.9%; Score 301; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 3.7e-144;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 TAGTCTTGGAAATTAACCTTGAACACCTTAAACCGCTTTAACCGAGCAACAGC 326
DB 301 TAGTCTTGGAAATTAACCTTGAACACCTTAAAGCTTTAACCGCTTTAACCGAGCAACAGC 242
QY 327 ATCAGAAGATAATGCTTTACAGAAATGGGAGCGATCTCTGAAATGCCCATTCAGTCGG 386
DB 241 ATCAGAAGATAATGCTTTACAGAAATGGGAGCGATCTCTGAAATGCCCATTCAGTCGG 182
QY 387 ACTTGATCGGTTCCACTCGGGATTTCAACATGCGCCAAAGAGCTTGCATACCTCT 446
DB 181 ACTTGATCGGTTCCACTCGGGATTTCAACATGCGCCAAAGAGCTTGCATACCTCT 122
QY 447 CCCGGTTTGAGAGCTGGACACCCAGGAGCGCGGTGTGTCCAGCTGATCAACCACTTGC 506

RESULT 11
AI028096 299 bp mRNA linear EST 28-AUG-1998
LOCUS
DEFINITION
Ow510.x1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGE:1650331 3', similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND
HAIRY-RELATED PROTEIN 1. ; mRNA sequence.

ACCESSION
AI028096
VERSION
AI028096.1 GI:3245405
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 299)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 331 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES
Location/Qualifiers
1..299
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1650331"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares parathyroid tumor NbHPA"
/notes="Organ: parathyroid gland; Vector: pT7T3D
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCGACCAATTTTTTTTTTTTTTTT
TTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

ORIGIN
Query Match 19.8%; Score 299; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 4e-143;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 TAGTCTTGGAAATTAACCTTGAACACCTTAAAGCTTTAACCGCTTTAACCGACAAACAG 325
```

Db		299	GTAGTCTTGGAAATTAACCTTTGAAACACTTTAAAAGCCTTTAACCGCCTTTAACCAGCAACAG	240
QY		326	CATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAATTCGCCCATTCAGTCC	385
Db		239	CATCAGAAGATAATTGCTTTACAGAATGGGAGCGATCTCTGAATTCGCCCATTCAGTCC	180
QY		386	GACTTTGGATCGTTCCACTCGGANNTTCAAACATGCAGCAAGAGTCTTGCAATACCTC	445
Db		179	GACTTTGGATCGTTCCACTCGGANNTTCAAACATGCAGCAAGAGTCTTGCAATACCTC	120
QY		446	TCCC GGTTTGAGAGCTGGACACACCAGGAGCCCGGTGTGTCCAGCTGATCAACCACTTG	505
Db		119	TCCC GGTTTGAGAGCTGGACACACCAGGAGCCCGGTGTGTCCAGCTGATCAACCACTTG	60
QY		506	CACGCCGTGGCCACCCAGTCTTTGCCACCCCGCCAGCTGTTGACTCAACAGGTCCTCT	564
Db		59	CACGCCGTGGCCACCCAGTCTTTGCCACCCCGCCAGCTGTTGACTCAACAGGTCCTCT	1
RESULT 12				
AI346780				
LOCUS		AI346780	451 bp mRNA linear	BST 02-FEB-1999
DEFINITION		qp53b04.xl NCI_CGAP C08 Homo sapiens cDNA clone IMAGE:1926703 3'		
			similar to contains TAR1.b1 MSRI repetitive element ; , mRNA sequence.	
ACCESSION		AI346780		
VERSION		AI346780.1	GI:4083986	
KEYWORDS		EST.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Euxheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		1. (bases 1 to 451)		
JOURNAL		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
COMMENT		National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. DNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnlnl.gov/bbrp/image/image.html Insert Length: 2359 Std Error: 0.00 Seq primer: -400P from Gibco. Location/Qualifiers 1. .451 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1926703" /tissue_type="adenocarcinoma" /lab_host="DHIOB" /clone_lib="NCI_CGAP C08" /note="Organ: colon; Vector: pTV73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."		
FEATURES				
source				
ORIGIN				
Query Match		19.1%	Score 288; DB 1; Length 451;	
Best Local Similarity		100.0%; Pred. No. 28-137;		
Matches 288; Conservative		0; Mismatches 0; Indels	0; Gaps	0;
QY		1224	CGCGCGCGCGCGCGCTTCCCTCGCTGCTCGCTGCTGTGCGGCTGTTGTCGCGCCCTCCCGAGAAG	1283

QY	1229	GC	CGCCGCGCGGTTCCCTCGCTGTCTCTCGGTGTGTTCGCGCCCTTCCCGAGAAAGCGGCG	1289							
Db	5	GC	CGCCGCGCGGTTCCCTCGCTGTCTCTCGGTGTGTTCGCGCCCTTCCCGAGAAAGCGGCG	64							
QY	1289	GC	CGCCGCGCGGACCTCTCTCGCGCAGCAGAGTGC	1348							
Db	65	GC	CGCGCGCGGACCTCTCTCGCGCAGAGTGC	124							
QY	1349	CACC	CGACGGCGCACCACCTTCCTTCGCGCGGCGCCCGGAGACCCGGAG	1408							
Db	125	CACC	CGACGGCGCACCACCTTCCTTCGCGCGGCGCCCGGAGACCCGGAG	184							
QY	1409	AGCT	CTCTCTCAGGAAGATCCCTCGCAGCAGGAAGAAAGCTCCCTCAATCCTTGCTGTC	1468							
Db	185	AGCT	CTCTCTCAGGAGATCCCTCGCAGCAGGAAGAAAGCTCCCTCAATCCTTGCTGTC	244							
QY	1469	CGAAG	AGCAGGTTCAAGCAGAGTGAGAGTTAAATACCT	1511							
Db	245	CGAAG	AGCAGGAGTTCAAGCAGAGTGAGAGTTAAATACCT	287							
RESULT 14											
LOCUS	CR615693	1109 bp mRNA linear HTC 21-JUL-2004									
DEFINITION	full-length cDNA clone CS0DE013YK10 of Placenta of Homo sapiens (human).										
ACCESSION	CR615693.1	GI:50496500									
VERSION	CR615693.1										
KEYWORDS	HTC; CNSLT cDNA.										
SOURCE	Homo sapiens (human)										
ORGANISM	Homo sapiens										
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.										
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.										
TITLE	Full-length cDNA libraries and normalization										
JOURNAL	Unpublished										
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope. 2 (bases 1 to 1109); Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime ends enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.										
FEATURES	Location/Qualifiers 1..1109 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DE013YK10" /tissue_type="placenta" /plasmid="pCMVSPORT_6"										
ORIGIN											
Query Match 18.3%; Score 276; DB 3; Length 1109;											
Best Local Similarity 100.0%; Pred. No. 3.5e-131;											
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
QY	1236	CCGCGTTCCCTCGCTGTTCGCTGTGTTCGCGCCCTTCCCGAGAAAGCGGCGCGCGCG	1295								
Db	1	CCGCGTTCCCTCGCTGTTCGCTGTGTTCGCGCCCTTCCCGAGAAAGCGGCGCGCGCG	60								
QY	1296	CCGCGACCTCTCTCGCGCAGAGTGC	1355								
Db	61	CCGCGACCTCTCTCGCGCAGAGTGC	120								

Qy	1356	ACGGCCGACACCCACTCTCCCTTCCTCGCGGCGCCCGCGAGACCGGGGAACCCGGAGAGCTCTG	1415
Db	121	ACGGCCGACACCCACTCTCCCTTCCTCGCGGCGCCCGCGAGACCGGGGAACCCGGAGAGCTCTG	180
Qy	1416	CTCAGGAAGATCCCTCGCAGCCAGGAAGAAAGCTCCCTGAATCTTTGGCTCCCGGAAGGA	1475
Db	181	CTCAGGAAGATCCCTCGCAGCCAGGAAGAAAGCTCCCTGAATCTTTGGCTCCCGGAAGGA	240
Qy	1476	CGGAGGTTCAAGCAGAGCTGAGAAGTTAAATACCTT	1511
Db	241	CGGAGGTTCAAGCAGAGCTGAGAAGTTAAATACCTT	276
RESULT 15			
RA982880/c			
LOCUS	AA983880	349 bp	mRNA linear EST 07-JUL-1998
DEFINITION	Op63C03.s1 Soares NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1581508 3' similar to TR:035779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED PROTEIN 1. ;, mRNA sequence.		
ACCESSION	AA983880		
KEYWORDS	AA983880.1 GI:3162405		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 349)		
TITLE	NCI-CPAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapsb-remail.nih.gov This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 612 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 348.		
FEATURES	Location/Qualifiers		
source	1. 349 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1581508" /lab_host="DH108" /clone_lib="Soares NFL_T_GBC_S1" /notes="Organ: pooled; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not 1; Site:2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CPAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 725408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaudo."		
ORIGIN			
Query Match	17.3%; Score 262; DB 1; Length 349;		
Best Local Similarity	100.0%; P: 0; Mismatches 0; Indels 0; Gaps 0;		
Matches 262; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	242	ACTCTGGGACATCTGGGAAAAGCTGTAGTCTTGGAATTAACCTTTGAAACACTTAAAAAGCT	301
Db	322	ACTCTGGGACATCTGGGAAAAGCTGTAGTCTTGGAATTAACCTTTGAAACACTTAAAAAGCT	263
Qy	302	TTAAACCGCTTAACCGAGCAACAGCATCAGAGATAATTCCTTTACAGAAATGGGAGCGGA	361
Db	262	TTAAACCGCTTAACCGAGCAACAGCATCAGAGATAATTCCTTTACAGAAATGGGAGCGGA	203

Qy	362	TCTCTGAAATCGCCATTTCAGTCCGACTTGGATGCGTTCGCTTCGCTTCGGGATTTCAAACATGC	421
Db	202	TCCTGAAATCGCCATTTCAGTCCGACTTGGATGCGTTCGCTTCGGGATTTCAAACATGC	143
Qy	422	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCGG	481
Db	142	GCCAAAGAAGTCTTGCAATACCTCTCCCGGTTTGAGAGCTGGACACCCAGGGAGCCGCGG	83
Qy	482	TGTGTCCAGCTGATCAACCACT	503
Db	82	TGTGTCCAGCTGATCAACCACT	61

Search completed: December 17, 2004, 16:09:38
 Job time : 4567.79 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	989	27.2	1109	3	CR615693 full-leng
2	683	18.8	715	4	BMS42244 AGENCOURT
3	663	18.2	892	5	BQ955258 AGENCOURT
4	657	18.0	901	1	AL542834 AL542834
5	626	17.2	665	5	BQ774398 UI-H-E21-
6	597	16.4	730	2	AW964159 EST376232
7	595	16.3	763	6	QA14694 UI-H-E20-
8	575	15.7	964	5	BQ219098 AGENCOURT
9	561	15.4	786	5	BQ447103 UI-H-EU1-
10	556	15.3	557	1	A1819798 wJ45a04.x
11	556	15.3	720	4	BG329136 HNC59-1-B
12	554	15.2	630	4	EG803221
13	539	14.8	757	6	CAS12575 UI-CF-FN0
14	523	14.4	523	1	A1672308 tY63f03.x
15	522	14.3	955	2	BF971232 602273249
16	518	14.2	653	5	BQ002547 UI-H-E11-
17	508	14.0	695	5	BX411208 EX411208
18	506	13.9	557	2	AW073334 wY97G05.x
19	506	13.9	647	7	CK430796 oJ5Aa07.y
20	502	13.8	736	5	BM978873 UI-CF-DU1
21	500	13.7	776	4	B1918832 603180555
22	497	13.7	638	5	BQ013516 UI-I-Bc1P
23	497	13.7	855	5	BX451442 EX451442
24	495	13.6	618	5	BU753502 UI-I-Bc1-

Db 363 CTGGGACATCTGGAGAAAGCTGTAGTCTTGGAAATTAACCTTTGAAACACCTTAAAGCTTTA 422
 Qy 432 ACCGCCCTTAACCGAGCAACAGCATCAGACGATTAATTCCTTTACAGAAATGGGAGCGATCT 491
 Db 423 ACCGCCCTTAACCGAGCAACAGCATCAGACGATTAATTCCTTTACAGAAATGGGAGCGATCT 482
 Qy 492 CTGAAATCGGCCATCTAGTCGGACTTGGATGCGTTCACACTCGGGATTTCAACATCGGCC 551
 Db 483 CTGAAATCGGCCATCTAGTCGGACTTGGATGCGTTCACACTCGGGATTTCAACATCGGCC 542
 Qy 552 AAGAAAGCTCTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGT 611
 Db 543 AAGAAAGCTCTGCAATACCTCTCCGGTTTGAGAGCTGGACACCCAGGGAGCGCGGTGT 602
 Qy 612 GTCAGCTGATCAACCACTTGACGCGCGTGGCCACCCAGTTCCTTGCCACCCCGCAGCTG 671
 Db 603 GTCAGCTGATCAACCACTTGACGCGCGTGGCCACCCAGTTCCTTGCCACCCCGCAGCTG 662
 Qy 672 TTGACTCAACAGGTCCTCTGAG 694
 Db 663 TTGACTCAACAGGTCCTCTGAG 685

RESULT 3
 BQ955258 892 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT 8732698 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6455656
 S', mRNA sequence.

ACCESSION BQ955258
 VERSION BQ955258.1 GI:22370736

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2634 row: k column: 17

High quality sequence stop: 684.

Location/Qualifiers

1. .892

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6455656"

/tissue type="epidermoid carcinoma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH_MGC_101"

/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:

XhoI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

ORIGIN

Query Match 18.2%; Score 663; DB 5; Length 892;

Best Local Similarity 100.0%; Pred. No. 2.9e-305;

Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1463 GCCGCAACCCGAGCACCAGCGGCGGCAACCCAGCTGCCCTTCGCGGGCCCCGGAGCC 1522
 Db 48 GCCGCAACCCGAGCACCAGCGGCGGCAACCCAGCTGCCCTTCGCGGGCCCCGGAGCC 107
 Qy 1523 GGGGAACCCGAGGAGCTCTGCTCAGGAAGATCCCTCGCAGCAGCAAGAGGAGCTCCCTG 1582
 Db 108 GGGGAACCCGAGGAGCTCTGCTCAGGAAGATCCCTCGCAGCAGCAAGAGGAGCTCCCTG 167
 Qy 1583 AATCCTTCGCTCCGAGGAGGATTCAGCAGAGAGTGAAGTAAATACCCCTTAAAGGAGAGTGA 1642
 Db 168 AATCCTTCGCTCCGAGGAGGATTCAGCAGAGAGTGAAGTAAATACCCCTTAAAG 227
 Qy 1643 GAGGTTCAAGCAGAGTGAAGTAAATACCCCTTAAAGGAGAGTGAAGTGA 1702
 Db 228 GAGGTTCAAGCAGAGTGAAGTAAATACCCCTTAAAGGAGAGTGAAGTGA 287
 Qy 1703 TAGATGCACGACGAGCATAAACAAGAACCAACAAACAGGTGTATGTATCATTCGGAGT 1762
 Db 288 TAGATGCACGACGAGCATAAACAAGAACCAACAAACAGGTGTATGTATCATTCGGAGT 347
 Qy 1763 TCCTGTTTGTCTCATCCGACACCCGACCCGACCTCCACACACTAACATCCCTTCTTCCCC 1822
 Db 348 TCCTGTTTGTCTCATCCGACACCCGACCCGACCTCCACACACTAACATCCCTTCTTCCCC 407
 Qy 1823 CCACGAGCTGTAAAGAGATCCTATGCGAAGAGACATGGCTCTTTTTTAAATCCCCCAAT 1882
 Db 408 CCACGAGCTGTAAAGAGATCCTATGCGAAGAGACATGGCTCTTTTTTAAATCCCCCAAT 467
 Qy 1883 AAATTTTGCCTTTTAGGCGCATGTTCCATTAICTCTTAAATTTGGAACCTTAATTCGAG 1942
 Db 468 AAATTTTGCCTTTTAGGCGCATGTTCCATTAICTCTTAAATTTGGAACCTTAATTCGAG 527
 Qy 1943 AGGAAGTGAAGAGGCTGTTCTGTGGCTGAGTAGGTGAACCCCGGGTAGGGGAAAGA 2002
 Db 528 AGGAAGTGAAGAGGCTGTTCTGTGGCTGAGTAGGTGAACCCCGGGTAGGGGAAAGA 587
 Qy 2003 TCTTAACACCTTTGACGCTTTTGGAGTTGACATGAACAGCAGGTAGTGTATGTATGAG 2062
 Db 588 TCTTAACACCTTTGACGCTTTTGGAGTTGACATGAACAGCAGGTAGTGTATGTATGAG 647
 Qy 2063 CTAGTCTCAAGCTGCCCTGCTGTTTAGAGGCGTTTCCACAAACAGATGAGGCTCT 2122
 Db 648 CTAGTCTCAAGCTGCCCTGCTGTTTAGAGGCGTTTCCACAAACAGATGAGGCTCT 707
 Qy 2123 TTT 2125
 Db 708 TTT 710

RESULT 4

AL542834

LOCUS

DEFINITION

AL542834 Homo sapiens PLACENTA

5-PRIME mRNA sequence.

ACCESSION

AL542834

VERSION

AL542834.3

GI:45718407

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 901)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

On Feb 15, 2001 this sequence version replaced gi:30548385.

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 6711.r

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?cs=CS05013BF05QPl&c=6711.r>.

FEATURES

Location/Qualifiers
1..901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS05013YK10"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 18.0%; Score 657; DB 1; Length 901;
Best Local Similarity 100.0%; Pred. No. 2.2e-302;
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1469 CCCGAGACCCGACGCGCGCACCCACCTGCTCCGCGGCGCGCGCGCGCGGAA 1528
Db 106 CCCGAGACCCGACGCGCGCACCCACCTGCTCCGCGGCGCGCGCGCGGAA 165
QY 1529 CCCGAGAGCTCTCTCAGGAAGATCCCTCGCAGCAGGAAGAGAGTCCCTGAATCCT 1588
Db 166 CCCGAGAGCTCTCTCAGGAAGATCCCTCGCAGCAGGAAGAGAGTCCCTGAATCCT 225
QY 1589 TGCCTCCGAGGACGAGGTTCAAGCAGAGTGAGAGTTAAATACCTTAAGAGGTT 1648
Db 226 TGCCTCCGAGGACGAGGTTCAAGCAGAGTGAGAGTTAAATACCTTAAGAGGTT 285
QY 1649 CACGACAGTGAGAGTTAAATACCTTAAGTCTTTAGGAGGAGTGAATAGATG 1708
Db 286 CACGACAGTGAGAGTTAAATACCTTAAGTCTTTAGGAGGAGTGAATAGATG 345
QY 1709 CACGACAGGATATAACAGAACACAAACAGGTGTTATGTGTACATTCGAGTTCCTGT 1768
Db 346 CACGACAGGATATAACAGAACACAAACAGGTGTTATGTGTACATTCGAGTTCCTGT 405
QY 1769 TTTGCTCATCCGACACACCCACCTCCACACTAACATCCCTTTCCGCCACCA 1828
Db 406 TTTGCTCATCCGACACACCCACCTCCACACTAACATCCCTTTCCGCCACCA 465
QY 1829 GCTGTAAAGATCCTATGCGAAGACACTGGCTCTTTTAAATCCCCCAATAATT 1888
Db 466 GCTGTAAAGATCCTATGCGAAGACACTGGCTCTTTTAAATCCCCCAATAATT 525
QY 1889 TGCCCCCTTTTAGGCCATGTCATTAATCTTTAAATTCGAACCTTAATTCGAGAGAG 1948
Db 526 TGCCCCCTTTTAGGCCATGTCATTAATCTTTAAATTCGAACCTTAATTCGAGAGAG 585
QY 1949 TACGAGGCTGTTCTGCTGAGCTAGCTGACACCCGCGGTAGGGAGAGATCTTAA 2008
Db 586 TACGAGGCTGTTCTGCTGAGCTAGCTGACACCCGCGGTAGGGAGAGATCTTAA 645
QY 2009 CACCTTTGACGTCTTTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAGTGT 2068
Db 646 CACCTTTGACGTCTTTGGAGTTGACATGGAACAGCAGGTAGTTGTTATGTAGAGTGT 705
QY 2069 CTCAGAGCTCCCTGCTGTTTAGGAGGCTTCCACACACAGATGAGGCTTTTT 2125
Db 706 CTCAGAGCTCCCTGCTGTTTAGGAGGCTTCCACACACAGATGAGGCTTTTT 762

RESULT 5

BO774398/c 665 bp mRNA linear EST 26-JUL-2002
LOCATION UI-H-EZ1-bca-h-10-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone

ACCESSION

BO774398

VERSION

BO774398.1 GI:21982874

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens (human)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 665)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of Orthopedics
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA sequence: 35-59, >Alu-richLow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..665
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EZ1-bca-h-10-0-UI"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ch2"
/note="Organ: Left Pelvis; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.
TAG_TISSUE=grade-2-chondrosarcoma
TAG_Lib=UI-H-EZ1
TAG_SEQ=ATCTAATATG"

ORIGIN

Query Match 17.2%; Score 626; DB 5; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.5e-287;
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2590 AACCAACACAAAAACAAAGTCCAAACTGATATATCTTATATCTTGTAAATTCAAAAGT 2649
Db 665 AACCAACACAAAAACAAAGTCCAAACTGATATATCTTATATCTTGTAAATTCAAAAGT 606
QY 2650 GAAAGAGAGCTTTAACTGGCCAGTTTGTGCAAAATGCTGTAAGATAGATGAAG 2709
Db 605 GAAAGAGAGCTTTAACTGGCCAGTTTGTGCAAAATGCTGTAAGATAGATGAAG 546
QY 2710 TCCTGTGAGGCTTCCTATCTCCAAAGTCTATGTTTTCTGGAGACCAACAGATACCA 2769
Db 545 TCCTGTGAGGCTTCCTATCTCCAAAGTCTATGTTTTCTGGAGACCAACAGATACCA 486
QY 2770 GATATCAAGAGAGCTTTTAAAGCTTTAAACCAAGCTTGTCTAGATATTTT 2829

QY 3003 AACTCTTAGCTCAAGGTGCTCTTATGACATAGTGTGTAATACATCCAAATTAATGAT 3062
 |||||
 Db 273 AACTCTTAGCTCAAGGTGCTCTTATGACATAGTGTGTAATACATCCAAATTAATGAT 214
 |||||
 QY 3063 GTCTGACATGCTATTTTGTAGGAGAAATATGCTGCTAATGATATTTTGTAGTAAATA 3122
 |||||
 Db 213 GTCTGACATGCTATTTTGTAGGAGAAATATGCTGCTAATGATATTTTGTAGTAAATA 154
 |||||
 QY 3123 TCTTTTGGGAGGATTTGCTGAAAGTTCACCTTTTGTACATGCTTATGCTTGTGTACA 3182
 |||||
 Db 153 TCTTTTGGGAGGATTTGCTGAAAGTTCACCTTTTGTACATGCTTATGCTTGTGTACA 94
 |||||
 QY 3183 AGCTTATGCTGCTTAAATTTTAAAAAAT 3215
 |||||
 Db 93 AGCTTATGCTGCTTAAATTTTAAAAAAT 61
 |||||

RESULT 9
 BQ447103/c
 LOCUS
 DEFINITION UI-H-EUI-bad-k-13-0-UI.s1 NCI CGAP Ctl Homo sapiens cDNA clone
 UI-H-EUI-bad-k-13-0-UI 3', mRNA sequence.
 BQ447103
 ACCESSION BQ447103.1 GI:21250215
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 786)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: ccapbs-x@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 28-52, >AT rich#Low complexity (matched complement)
 632-675, >(CAAAA)n#Simple_repeat (matched complement)
 Seq primer: M13 FORWARD
 POLYA-Yes.

Location/Qualifiers
 1. 786
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-EUI-bad-k-13-0-UI"
 /tissue_type="Osteoarthritic Cartilage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Ctl"
 /notes="Organ: Knee; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site: 1: Ecor I; Site 2: Not I;
 NCI CGAP Ctl is a normalized cDNA library containing the
 following tissue(s): Osteoarthritic Cartilage The library
 was constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into p773-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dr)18 tail. The sequence tag for this library is
 TGATCACGCT.
 TAG_TISSUE=osteoarthritic cartilage

FEATURES

source
 1. 786
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-EUI-bad-k-13-0-UI"
 /tissue_type="Osteoarthritic Cartilage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Ctl"
 /notes="Organ: Knee; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site: 1: Ecor I; Site 2: Not I;
 NCI CGAP Ctl is a normalized cDNA library containing the
 following tissue(s): Osteoarthritic Cartilage The library
 was constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into p773-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dr)18 tail. The sequence tag for this library is
 TGATCACGCT.
 TAG_TISSUE=osteoarthritic cartilage

TAG LIB=UI-H-EUI
 TAG_SEQ=TGATCACGCT"

ORIGIN

Query Match 15.4%; Score 561; DB 5; Length 786;
 Best Local Similarity 99.9%; Pred. No. 1.8e-256;
 Matches 681; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 2534 TTTTGTAAATCTTACTGTAATAAAGTGCACCTGAACTTTCAAAACAAAAACAACA 2593
 |||||
 Db 713 TTTTGTAAATCTTACTGTAATAAAGTGCACCTGAACTTTCAAAACAAAAACAACA 654
 |||||
 QY 2594 ACAACAAAAACAAGTCCAACTGATATATCTCTATATTTCTGTTAAATTTCAAAGTGAAC 2653
 |||||
 Db 653 ACAACAAAAACAAGTCCAACTGATATATCTCTATATTTCTGTTAAATTTCAAAGTGAAC 594
 |||||
 QY 2654 GAAAGCATTTAACTGGCCAGTTTTCATTGCAATGCTGTAAGATATATAGATGAAGTCT 2713
 |||||
 Db 593 GAAAGCATTTAACTGGCCAGTTTTCATTGCAATGCTGTAAGATATATAGATGAAGTCT 534
 |||||
 QY 2714 GTGAGGCTTCTCTATCTCCAGTCTATGATATTTCTGGAGACCAACCCAGATACCAGATA 2773
 |||||
 Db 533 GTGAGGCTTCTCTATCTCCAGTCTATGATATTTCT-GAGACCAACCCAGATACCAGATA 475
 |||||
 QY 2774 ATCACAAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGTCTAGATATTTTACT 2833
 |||||
 Db 474 ATCACAAAAGAAAGCTTTTAAATAGGCTTAAACCAAGACCTTGTCTAGATATTTTACT 415
 |||||
 QY 2834 TTGTTGCCAAGTAGCAGTGTGAGAAATCTCACITGGATGTTATGTAAGGGGTGAGACAC 2893
 |||||
 Db 414 TTGTTGCCAAGTAGCAGTGTGAGAAATCTCACITGGATGTTATGTAAGGGGTGAGACAC 355
 |||||
 QY 2894 AACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCTGTCAGTTTGTGCAATTTGC 2953
 |||||
 Db 354 AACAGTCTGACTATGAGTGAGGAAATATCTGGGTCTTTTCTGTCAGTTTGTGCAATTTGC 295
 |||||
 QY 2954 TCTCTCTGTGTCTACTGTTTGCCTCAAAAGCTGTGTTTAAACAAAGCTTAAACTCTTAGCC 3013
 |||||
 Db 294 TCTCTCTGTGTCTACTGTTTGCCTCAAAAGCTGTGTTTAAACAAAGCTTAAACTCTTAGCC 235
 |||||
 QY 3014 TACAAGTGGCTCTTATGATACATAGTGTGTTAATACATCAATTAATGATGTGACATGC 3073
 |||||
 Db 234 TACAAGTGGCTCTTATGATACATAGTGTGTTAATACATCAATTAATGATGTGACATGC 175
 |||||
 QY 3074 TATTTTCTGAGGAGAAATATGCTAATGATATTTTGAAGTTAAATATCTTTTGGGGA 3133
 |||||
 Db 174 TATTTTCTGAGGAGAAATATGCTAATGATATTTTGAAGTTAAATATCTTTTGGGGA 115
 |||||
 QY 3134 GGATTTGCTGAAAAGTTCACCTTTTGTGTTACAAATGCTTATGCTTGTGTAAGCTTATGCTG 3193
 |||||
 Db 114 GGATTTGCTGAAAAGTTCACCTTTTGTGTTACAAATGCTTATGCTTGTGTAAGCTTATGCTG 55
 |||||
 QY 3194 TCTTAAATATTTTAAAAAAT 3215
 |||||
 Db 54 TCTTAAATATTTTAAAAAAT 33
 |||||

RESULT 10

AI819798 557 bp mRNA linear EST 21-DEC-1999
 wj4s04.xl NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2405742 3'
 similar to TR:O35779 O35779 ENHANCER-OF-SPLIT AND HAIRY-RELATED
 PROTEIN 1. ; mRNA sequence.
 ACCESSION AI819798
 VERSION AI819798
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 557)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csapsb@mail.nih.gov
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/brp/image/image.html
Insert Length: 340 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 441.

FEATURES
source
1..557
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2405742"
/tissue_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"
/dev_stage="adult"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lul9"
note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 15.3%; Score 556; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.6e-254; Indels 0; Gaps 0;
Matches 556; Conservative 0; Mismatches 0;

QY 135 ATGACGAGGAATTCCTCATTTCCAGAGAGAGAGAGTACTGGAACATAGAGATTTTATA 194
Db 557 ATGACGAGGAATTCCTCATTTCCAGAGAGAGAGTACTGGAACATAGAGATTTTATA 498
QY 195 GGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAAGGAGCATGAAACGAGAGCAG 254
Db 497 GGACTGGACTATTCCTCTTTGTATATGTGTAACCCAAAAGGAGCATGAAACGAGAGCAG 438
QY 255 ACCAAGGATACCTACAAATACCGACAGATTATAGAAAGAAAGAAAGAGACCGAATT 314
Db 437 ACCAAGGATACCTACAAATACCGACAGATTATAGAAAGAAAGAGACCGAATT 378
QY 315 AATGAATGCAATGCTGAGCTGAAAGATTTACTGCTGAAACATCTGAAATGCAACTCTG 374
Db 377 AATGAATGCAATGCTGAGCTGAAAGATTTACTGCTGAAACATCTGAAATGCAACTCTG 318
QY 375 GGACATCTGGGAAGCTGTAGCTTGGAAATTAACCTTTGAACACATTAAGGCTTTAACC 434
Db 317 GGACATCTGGGAAGCTGTAGCTTGGAAATTAACCTTTGAACACATTAAGGCTTTAACC 258
QY 435 GCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGGAGCGATCTCTG 494
Db 257 GCCTTAACCGAGCAACAGCATCAGAAGATAATTGCTTTACAGAATGGGGAGCGATCTCTG 198
QY 495 AAATCGGCCATTCAGTCGACCTGGATCGCTTCACCTCGGGATTTGAAACATGGCCAAA 554
Db 197 AAATCGGCCATTCAGTCGACCTGGATCGCTTCACCTCGGGATTTGAAACATGGCCAAA 138
QY 555 GAAGCTTTGCAATACCTCTCCCGTTTGGAGAGCTGGACACCCAGGAGCGCGGTGTGTC 614
Db 137 GAAGCTTTGCAATACCTCTCCCGTTTGGAGAGCTGGACACCCAGGAGCGCGGTGTGTC 78

QY 615 CAGCTGATCAACCACTTGCACGCGTGGCCACCCAGTCTTGTGCCACCCCGAGCTTTG 674
Db 77 CAGCTGATCAACCACTTGCACGCGTGGCCACCCAGTCTTGTGCCACCCCGAGCTTTG 18
QY 675 ACTCAACAGGTCCTC 690
Db 17 ACTCAACAGGTCCTC 2

RESULT 11
BG29136
LOCUS
DEFINITION HNC59-1-B9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
sequence.
ACCESSION BG29136
VERSION BG29136.1 GI:14323659
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 720)
AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J., Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and Lark,M.W.
TITLE Identification and initial characterization of 5000 expressed sequence tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries
JOURNAL Osteoarthritis. Cartil. 9 (7), 641-653 (2001)
MEDLINE 21482651
PUBMED 11597177
COMMENT Contact: Sanjay Kumar
UW2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@gsk.com
Seq primer: 77
Location/Qualifiers
1..720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/clone_lib="HNC (Human Normal Cartilage)"
/note="vector: pSPORT 1; Site_1: SalI; Site_2: NotI; Directional"

ORIGIN
Query Match 15.3%; Score 556; DB 4; Length 720;
Best Local Similarity 99.8%; Pred. No. 4.5e-254; Indels 0; Gaps 0;
Matches 606; Conservative 0; Mismatches 1;

QY 2587 AACCAACAACAACAAACAAAGTCCAACTGATATATCTGTAAGATATAGAATG 2646
Db 89 AACCAACAACAACAAACAAAGTCCAACTGATATATCTGTAAGATATAGAATG 148
QY 2647 AGTGAACGAAGCATTTAACTGGCCAGTTTGAATGCAATGCTTAAGATATAGAATG 2706
Db 149 AGTGAACGAAGCATTTAACTGGCCAGTTTGAATGCAATGCTTAAGATATAGAATG 208
QY 2707 AAGTCTCTGGAGGCTTCTCTATCTCCAAAGTCTATGTAATTTCTGGAGACCAACAGATA 2766
Db 209 AAGTCTCTGGAGGCTTCTCTATCTCCAAAGTCTATGTAATTTCTGGAGACCAACAGATA 268
QY 2767 CCAGATAATCAACAAGAAGCTTTTAAAGCTTTAAACCAAGACCTTCTGTAGATAT 2826
Db 269 CCAGATAATCAACAAGAAGCTTTTAAAGCTTTAAACCAAGACCTTCTGTAGATAT 328
QY 2827 TTTTAGTTGTTGCCAAGGTAGCACTGTGAGAAATCTCACTTGGATGTTATGTAAGGGT 2886

```

Db      329  TTTTAGTTTGTGCAAGGTAGCACTGTGAGAAATCTCACTTGATGTTATGTAAGGGGT 388
Qy      2887  GAGACACACAGTCTGACTACTGAGTGGAGAAATATCTGGGTCTTTTTCGTCAGTTGGTG 2946
Db      389  GAGACACACAGTCTGACTACTGAGTGGAGAAATATCTGGGTCTTTTTCGTCAGTTGGTG 448
Qy      2947  CATTTGCTGCTGCTGTTGCTACTGTTTGCCTCAACGCTGTTTAAACACGTTAACT 3006
Db      449  CATTTGCTGCTGCTGTTGCTACTGTTTGCCTCAACGCTGTTTAAACACGTTAACT 508
Qy      3007  CTTAGCCTCAAGGTGGCTCTTATGTACATAGTTGTTTAAACATCCCAATTAATGATGCT 3066
Db      509  CTTAGCCTCAAGGTGGCTCTTATGTACATAGTTGTTTAAACATCCCAATTAATGATGCT 568
Qy      3067  GACATGCTATTTTGTAGGAGAAATATGTGCTAATGATATTTGAGTTAAATATCTT 3126
Db      569  GACATGCTATTTTGTAGGAGAAATATGTGCTAATGATATTTGAGTTAAATATCTT 628
Qy      3127  TTGGGGAGGATTTCCTGAAAGTTGCACTTTTGTACAAGTCTTATGCTTGGTACAAGCT 3186
Db      629  TTGNGAGGATTTCCTGAAAGTTGCACTTTTGTACAAGTCTTATGCTTGGTACAAGCT 688
Qy      3187  TATGCTG 3193
Db      689  TATGCTG 695

```

```

RESULT 12
BG680321
LOCUS
DEFINITION 602629213P1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754019 5',
mRNA sequence.

```

```

ACCESSION BG680321
VERSION BG680321.1 GI:13911718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 630)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10614 rcw: j column: 04
High quality sequence stop: 627.
Location/Qualifiers
1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754019"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

```

FEATURES
source
1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754019"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

```

ORIGIN
Query Match 15.2%; Score 554; DB 4; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.1e-253;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2123  TTTTAGAATGTAATTTACTCTTCACTATTTTCTTAATGTTCAAGCTTTCTAAAGGCATATA 2182
Db      48  TTTTAGAATGTAATTTACTCTTCACTATTTTCTTAATGTTCAAGCTTTCTAAAGGCATATA 107
Qy      2183  TTTTTCACAAAGAGTGCAGATGCTCTCAGTTGCAACCTATTCTCAAGTGGTTTAAA 2242
Db      108  TTTTTCACAAAGAGTGCAGATGCTCTCAGTTGCAACCTATTCTCAAGTGGTTTAAA 167
Qy      2243  TCGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAAC 2302
Db      168  TCGTATCTCTTAGTAACCTTGCACTCGTTAAAGAAACACGGAGCTGGGCCATCGTCAGAAC 227
Qy      2303  TAAAGTCAGGAGGAGATGGATGAGAAGGCCAGATCAATCTCTAGTACATTTGCTAAACAC 2362
Db      228  TAAAGTCAGGAGGAGATGGATGAGAAGGCCAGATCAATCTCTAGTACATTTGCTAAACAC 287
Qy      2363  TTTTATTGAGAAATGACCATGAATTAATGGACTCATCTTAATTTCTTCAAGTCCATATA 2422
Db      288  TTTTATTGAGAAATGACCATGAATTAATGGACTCATCTTAATTTCTTCAAGTCCATATA 347
Qy      2423  TAGATAGATATCTATCTGTACAGATTTCTATTATTCATAGATAGATGATCTATACATACA 2482
Db      348  TAGATAGATATCTATCTGTACAGATTTCTATTATTCATAGATAGATGATCTATACATACA 407
Qy      2483  CATCTCAAGTGCATCTATTCCACATCTCATTAATCCATCATGTTCTCTAAATTTTGTAAAT 2542
Db      408  CATCTCAAGTGCATCTATTCCACATCTCATTAATCCATCATGTTCTCTAAATTTTGTAAAT 467
Qy      2543  CTTACTGTAAAAAAGTGCACCTGAATCAAAACAAAAACAAAAACAAACAAACAAAA 2602
Db      468  CTTACTGTAAAAAAGTGCACCTGAATCAAAACAAAAACAAAAACAAACAAACAAAA 527
Qy      2603  AACAAAGTCCAACTCATATATCTCTTAAATTCGTTAAATTCGAAAGTGAACGAAAGCATT 2662
Db      528  AACAAAGTCCAACTCATATATCTCTTAAATTCGTTAAATTCGAAAGTGAACGAAAGCATT 587
Qy      2663  TAACTGGCCAGTTT 2676
Db      588  TAACTGGCCAGTTT 601

```

```

RESULT 13
CA312575/c
LOCUS
DEFINITION 757 bp mRNA linear EST 04-NOV-2002
UI-CF-FNO-afk-h-06-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
UI-CF-FNO-afk-h-06-0-UI 3', mRNA sequence.

```

```

ACCESSION CA312575
VERSION CA312575.1 GI:24530673
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research

```


Db	301	CTTGGGCGCGCAGCCCGCAGCACCCCGCAGCGCCGCGCACCCACCTGCGCTTCGCCGCGGCGCC	360		
Qy	1515	CGGAGCGCGGGAAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAA	1574		
Db	361	CGCGAGCGCGGAAACCCGGAGAGCTCTGCTCAGGAAGATCCCTCGCAGCCAGGAAGGAA	420		
Qy	1575	GCTCCCTGAATCCTTGGTCCGCGAGCAGCGAGGTTCAAGCAGAGTCGAGAGTTAAATA	1634		
Db	421	GCTCCCTGAATCCTTGGTCCGCGAGCAGCGAGGTTCAAGCAGAGTCGAGAGTTAAATA	480		
Qy	1635	CCCTTAAGGAGGTTCAAGCAGAGTCGAGAGTTAAATAACCCCTT	1677		
Db	481	CCCTTAAGGAGGTTCAAGCAGAGTCGAGAGTTAAATAACCCCTT	523		
RESULT 15					
LOCUS	BF971232	602273249f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361004 5',	EST 22-JAN-2001		
DEFINITION	mRNA sequence.				
ACCESSION	BF971232				
VERSION	BF971232.1	GI:12338447			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-x@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10003 Row: b column: 13 High quality sequence stop: 638. Location/Qualifiers 1. 955 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4361004" /tissue_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC 84" /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr_ primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."				
FEATURES					
source					
ORIGIN					
Query Match	14.3%; Score 522; DB 2; Length 955;				
Best Local Similarity	100.0%; Pred. No. 8e-238;				
Matches 522;	Conservative	0; Mismatches	0; Indels 0; Gaps 0;		
Qy	2566	GAACCTTCAAAACAAAAACAAAAACAAACAAACAAACAAAGTCCTATATATATATCC	2625		
Db	1	GAACCTTCAAAACAAAAACAAAAACAAACAAACAAACAAAGTCCTATATATATATCC	60		
Qy	2626	TATATTCGTGTAAATTCAAAAGTGAACGAAAGCATTTAACTGCCAGTTTTTCATTGCAA	2685		
Db	61	TATATTCGTGTAAATTCAAAAGTGAACGAAAGCATTTAACTGCCAGTTTTTCATTGCAA	120		
Qy	2686	ATGCTGTAAGATATAGAATGAAGTCTGTGAGCGCTTCTATCTCAAGTCATATGATT	2745		
Db	121	ATGCTGTAAGATATAGAATGAAGTCTGTGAGCGCTTCTATCTCAAGTCATATGATT	180		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: December 15, 2004, 19:33:10 ; Search time 196 Seconds
(without alignments)
1420.823 Million cell updates/sec

Title: US-10-078-650-12

Perfect score: 2517

Sequence: 1 WDEGIPHLQERLLEHRDFI.....PCNPSSAQEDPSQPGKEAP 484

Scoring table: BLOSUM62

Searched: 1825181 seqs, 575374546 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_spot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	2498	99.2	482	1 BHB3_HUMAN	Q9C0J9 homo sapien
2	2494	99.1	482	2 Q8TAT1	Q8tat1 homo sapien
3	2328.5	92.5	513	2 Q7YQC9	Q7Yqc9 canis famli
4	1634	64.9	410	1 BHB3_MOUSE	Q99qv5 mus musculu
5	1634	64.9	410	2 BAI18970	Bai18970 mus muscu
6	1624	64.5	410	1 BHB3_RAT	Q35779 rattus norv
7	905.5	36.0	193	2 Q8GCH4	Q8gch4 mus musculu
8	685.5	27.2	411	1 BHB2_MOUSE	Q35185 mus musculu
9	685.5	27.2	411	2 BAC33281	Bac33281 mus muscu
10	685.5	27.2	411	2 BAC41056	Bac41056 mus muscu
11	684.5	27.2	403	2 Q6NY50	Q6ny50 brachydania
12	684.5	27.2	403	2 AAH66738	Aah66738 brachydani
13	682.5	27.1	411	1 BHB2_RAT	Q35780 rattus norv
14	679.5	27.0	411	2 Q76JQ4	Q76jq4 rattus norv
15	679.5	27.0	411	2 BAB01388	Bab01388 rattus no
16	669	26.6	412	1 BHB2_HUMAN	Q14503 homo sapien
17	669	26.6	412	2 CAG33202	Cag33202 homo sapi
18	589.5	23.4	409	2 Q6GNE9	Q6gne9 xenopus lae
19	242	9.6	786	2 Q7QJD2	Q7qjd2 anopheles g
20	230.5	9.2	610	2 Q9U450	Q9u450 drosophila
21	230.5	9.2	598	2 Q9VGZ5	Q9vgz5 drosophila
22	220.5	8.8	333	2 Q8CD44	Q8cd44 mus musculu
23	220.5	8.8	339	2 Q9QUS4	Q9qus4 mus musculu
24	217.5	8.6	317	2 Q8AXV6	Q8axv6 brachydania
25	215.5	8.6	769	2 Q9LLJ1	Q9llj1 chlamydomon
26	215	8.5	3288	2 Q7T5D9	Q7t5d9 cercopithe
27	209.5	8.3	460	2 Q9GZP7	Q9gzf7 caenorhabdi
28	207.5	8.2	3326	2 Q7T591	Q7t591 cercopithe
29	207	8.2	337	2 Q9UBP5	Q9ubp5 homo sapien
30	207	8.2	668	2 Q8MXU8	Q8mxu8 caenorhabdi
31	205.5	8.2	1953	2 Q9BIT7	Q9bit7 nephrila ina

RESULT 1

BHB3_HUMAN

ID BHB3_HUMAN STANDARD; PRT; 482 AA.

AC Q9C0J9; (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially

DE expressed in chondrocytes protein 2) (hDEC2) (Enhancer-of-split and

DE hairy-related protein 1) (SHARP-1).

GN Name=BHLHB3; Synonyms=DEC2, SHARP1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCSI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21092582; PubMed=11162494; DOI=10.1006/bbrc.2000.4133;

RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,

RA Yoshida E., Suadita K., Matsuda Y., Kato Y.;

RT "Molecular cloning and characterization of DEC2, a new member of basic

RT helix-loop-helix proteins.";

RL Biochem. Biophys. Res. Commun. 280:164-171(2001).

RN [2]

RN FUNCTION.

RX MEDLINE=21226716; PubMed=11278948; DOI=10.1074/jbc.M011619200;

RA Garriga-Canut M., Roopra A., Buckley N.J.;

RT "The basic helix-loop-helix protein, SHARP-1, represses transcription

RT by a histone deacetylase-dependent and histone deacetylase-independent

RT mechanism.";

RL J. Biol. Chem. 276:14821-14828(2001).

CC -!- FUNCTION: May be a transcriptional repressor that represses both

CC basal and activated transcription.

CC -!- SUBUNIT: Homodimerize.

CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).

CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle and brain,

CC moderately expressed in pancreas and heart, weakly expressed in

CC placenta, lung, liver and kidney.

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC -!- SIMILARITY: Contains 1 orange domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB044088; BAB21502.1; -

DR PIR; JC7583; JC7583.

DR Genew; HGNC:16617; BHLHB3.

DR MIM; 606200; -

DR GO; GO:0005634; C:nucleus; NAS.

DR


```
QY 421 LSPPEKAGAAAATLLPHEVAPLGAHPHGHGTHLPAGPREPCNPRESSAQEDPSQPG 480
Db 419 LSPPEKAGAAAATLLPHEVAPLGAHPHGHGTHLPAGPREPCNPRESSAQEDPSQPG 478
QY 481 KEAP 484
Db 479 KEAP 482

RESULT 3
QYQC9 PRELIMINARY; PRT; 513 AA.
AC QYQC9;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE SHARPI protein.
GN Name=SHARPI;
OS Canis familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=22791403; PubMed=12909371;
RA Kukekova A.V., Aguirre G.D., Acland G.M.;
RT "Cloning and characterization of canine SHARPI and its evaluation as a
RT positional candidate for canine early retinal degeneration (erd).";
RL Gene 312:335-343 (2003).
DR EMBL; AY204567; AAP12464.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_Orange; 1.
DR SMART; SM00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00888; HLH; 1.
SQ SEQUENCE 513 AA; 52942 MW; CA7FC23C2E03DCBA CRC64;

Query Match 92.5%; Score 2328.5; DB 2; Length 513;
Best Local Similarity 90.3%; Pred. No. 4.8e-113;
Matches 465; Conservative 4; Mismatches 11; Indels 35; Gaps 6;

QY 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 60
Db 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 58

QY 61 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLLKHLKALTALTEQOQKIIALQNGER 120
Db 59 RINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLLKHLKALTALTEQOQKIIALQNGER 118

QY 121 SLKSPQSOLDLAFHSGFTCAKEVLQYLSRFSWTPREPRVCVQLNHLHVAVTOPLPTPQ 180
Db 119 SLKSPQSOLDLAFHSGFTCAKEVLQYLSRFSWTPREPRVCVQLNHLHVAVTOPLPTPQ 178

QY 181 LITQOVPLSKGTGAPSA---AGSAAPCLERAGQKLEPLAYCVPIQRTQPSAEIAEND 237
Db 179 LITQOVPLSKGAPSAAPAGSAAPCLERAGQKLEFLAHCVPIQRTQPSAEIAEND 238

QY 238 TTDSGYGGEAEARDREKKGAGASRVTIKQEPFGEDSPAPKMKLDSRGSGGGG-- 295
Db 239 TTDSGYGGEAEARDRKGGKSGAGRVTIKQEPFGEDSPAPKMKLDSRGSGGGGGG 298

QY 296 -----GGGAAAAAALLGPPAAAAALLPDAALLSSILVAFGGGGAP 338
Db 299 LGGGGGGGLGGGGGLGGGAAAAAALLGPPAAAAALLPDAALLSSILVAFGGGGAP 358

QY 339 FPQAAAAAPFLPCFLSPSAAAAYVQPLDKSGLEKLYLPAAAAAPFLLYPCIPAPA 398
Db 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 60
Db 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 58
```

```
Db 359 FAQPAAMAAAPCLPFYFLSPSAAAAYVQPLDKSGLEKLYLPAAAAAPFLLYPCIPAP- 417
QY 399 AAAAAAAMAAAPCLSSVLSPPPEKAGAAAATLLPHEVAPLGA-----PFPQ 449
Db 418 AAAAAAAMAAAPCLSSVLSPPPEKAGAAAATLLPHEVAPLGA-----PFPQ 477
QY 450 -HPHGRTHLPAGPREPCNPRESSAQEDPSQPGKEA 483
Db 478 PPHGRTHLPAGPREPCNPRESSAQEDPSQPGKEA 512

RESULT 4
BHE3 MOUSE
ID BHE3 MOUSE STANDARD; PRT; 410 AA.
AC Q99FV5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 3 (bHLHB3) (Differentially
DE expressed in chondrocytes protein 2) (mDEC2).
GN Name=Bhlhb3; Synonyms=Dec2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21095582; PubMed=1162494; DOI=10.1006/bbrc.2000.4133;
RA Fujimoto K., Shen M., Noshiro M., Matsubara K., Shingu S., Honda K.,
RA Yoshida E., Suadrita K., Matsuda Y., Kato Y.;
RT "Molecular cloning and characterization of DEC2, a new member of basic
RT helix-loop-helix proteins.";
RL Biochem. Biophys. Res. Commun. 280:164-171 (2001).
CC -! FUNCTION: May be a transcriptional repressor that represses both
CC basal and activated transcription.
CC -! SUBUNIT: Homodimerize.
CC -! SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -! TISSUE SPECIFICITY: Expressed in skeletal muscle, brain and lung.
CC -! SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.
CC -! SIMILARITY: Contains 1 orange domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB044090; BAB21503.1; -.
DR PIR; JC7584; JC7584.
DR MGI; MGI:1930704; Bhlhb3.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003650; Orange.
DR Pfam; PF07527; Hairy_Orange; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00511; ORANGE; 1.
DR PROSITE; PS00888; HLH; 1.
DR PROSITE; PS00511; ORANGE; 1.
KW DNA-binding; Nuclear protein; Repressor; Transcription regulation.
FT DNA_BIND 45 57 Basic motif.
FT DOMAIN 58 100 Helix-loop-helix motif.
FT DOMAIN 129 175 Orange.
FT DOMAIN 321 373 Ala/Gly-rich.
SQ SEQUENCE 410 AA; 43946 MW; 40A87281B08E233D CRC64;

Query Match 64.9%; Score 1634; DB 1; Length 410;
Best Local Similarity 72.0%; Pred. No. 3.9e-77;
Matches 350; Conservative 21; Mismatches 37; Indels 78; Gaps 16;

QY 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 60
Db 1 MDEGPHLQEROLLEHRDFIGLDYSLYCKPKRSMKRDYTKVDTYKLPRLIEKKRD 58
```


CC frameshift in position 249.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF009329; AAB63586.1; ALT_FRAME.

CC RGD; 70900; Bhlhb3.

CC InterPro; IPR001092; HLH basic.

CC Pfam; PF07527; Hairy orange; 1.

CC Pfam; PF00010; HLH; 1.

CC SMART; SM00353; HLH; 1.

CC SMART; SM00511; ORANGE; 1.

CC PROSITE; PS00888; HLH; 1.

CC DNA BIND 45 57

CC DNA-binding; Nuclear protein; Repressor; Transcription regulation.

CC FT DOMAIN 58 100

CC FT DOMAIN 129 175

CC FT DOMAIN 321 344

CC FT DOMAIN 410 AA; 43917 MW; 829705CA3A013127 CRC64;

CC SEQUENCE 410 AA; 43917 MW; 829705CA3A013127 CRC64;

CC -----

CC Query Match 64.5%; Score 1624; DB 1; Length 410;

CC Best Local Similarity 71.6%; Pred. No. 1.3e-76;

CC Matches 348; Conservative 23; Mismatches 37; Indels 78; Gaps 16;

CC

CC 1 MDEGPHLQERQLLEHRDFIGLDSLYMCKPKSRMKRDTKVSDTYKLPRLIEKKRD 60

CC 1 MDEGPHLQERQLLEHRDFIGLDSLYMCKPKSRMKRDTKVSDTYKLPRLIEKKRD 58

CC 61 RINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIALLONGER 120

CC 59 RINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIALLONGER 118

CC 121 SLKSPIQSDLDAPHSGFQTCACEVQLYLSRFESWTPREPRCVQLINHLHVAATQFLPTPQ 180

CC 119 SLKSPVQADLDAPHSGFQTCACEVQLYLSRFESWTPREPRCAQLVSHLHAVAT-----Q 172

CC 181 LITQVPLSKGTG-AFSAAGS 200

CC 173 LITQVPLSKGTG-AFSAAGS 193

CC 240 TDSGYGGEAEARPDREKKGAGASVTIKQEPGDSAPKRMKLDNRGGSGGPGGGA 299

CC 224 TDSGYGGEAE-----QGRAVVKQEPGDSAPKRMKLEAG----- 260

CC 300 AAAAAALQPDPAALALRPAALSSLVAFGGGGAPFPQAAAAAPFCLPFCFLSPS 359

CC 261 -----ALLGFEP-----ALLGSLVAL--GGGAPFAQF--AAAPFCLPFYLLSPS 300

CC 360 AAAAAVQPLDLSGLKLYLPAAAAAPPLYPGPAPAAAAAALAAAAAAPPCLSS 419

CC 301 -AAAVVQPLDLSGLKLYLP-AAAAPFLLYEGIP-----AAAAAALAAAAAPPCLSS 350

CC 420 VLSPPPEKAGAAA-ATLLPHEVAPLGAHPQPHQTHLPFAGPREPCNPRESSAQEDPSQ 478

CC 351 VLSPPPEKAGSAGAPFLAHEVAPPGSLRPOHAHSTHLP-----VNPRESS-QEDATQ 404

CC 479 PGKEAP 484

CC 405 PAKDAP 410

CC

CC RESULT 7

CC Q8CGH4 PRELIMINARY; PRT; 193 AA.

CC AC Q8CGH4;

CC DT 01-MAR-2003 (TReMBLrel. 23, Created)

CC DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Bhlhb3 protein (Fragment).

GN Name=Bhlhb3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary Gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Touchman J.W., Green E.D., Dickinson G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson G.G.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RL and mouse cDNA sequences."

RP Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary Gland;

RX Strausberg R.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC037742; AAK37742.1; -

DR MGD; MGI:1930704; Bhlhb3.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001092; HLH basic.

DR InterPro; IPR003650; Orange.

DR Pfam; PF07527; Hairy orange; 1.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR SMART; SM00511; ORANGE; 1.

DR PROSITE; PS00888; HLH; 1.

FT NON TER 193

SQ SEQUENCE 193 AA; 21929 MW; 854890594639A422 CRC64;

Query Match 36.0%; Score 905.5; DB 2; Length 193;

Best Local Similarity 88.1%; Pred. No. 1.1e-39;

Matches 177; Conservative 10; Mismatches 5; Indels 9; Gaps 3;

Qy 1 MDEGPHLQERQLLEHRDFIGLDSLYMCKPKSRMKRDTKVSDTYKLPRLIEKKRD 60

Db 1 MDEGPHLQERQLLEHRDFIGLDSLYMCKPKSRMKRDTKVSDTYKLPRLIEKKRD 58

Qy 61 RINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIALLONGER 120

Db 59 RINECIAQLKDLLPEHLKLTTLGHLKAVVLETLKHLKALTALTEQOQKIALLONGER 118

Qy 121 SLKSPIQSDLDAPHSGFQTCACEVQLYLSRFESWTPREPRCVQLINHLHVAATQFLPTPQ 180

Db 119 SLKSPVQADLDAPHSGFQTCACEVQLYLSRFESWTPREPRCAQLVSHLHAVAT-----Q 172

Qy 181 LITQVPLSKGTG-AFSAAGS 200

Db 173 LITQVPLSKGTG-AFSAAGS 193

RESULT 8

BHB2_MOUSE


```

RESULT 9
BAC33281
ID BAC33281 PRELIMINARY; PRT; 411 AA.
AC BAC33281;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DE 16 days embryo head cDNA, Riken full-length enriched library,
DE clone:130042M06 product:basic helix-loop-helix domain containing,
DE class B2, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The RIKEN Consortium.
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA "RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA "RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771 (2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,

```

```

RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK048242; BAC33281.1; -
SQ SEQUENCE 411 AA; 45360 MW; B392893CD49292BC CRC64;
Query Match 27.2%; Score 685.5; DB 2; Length 411;
Best Local Similarity 43.1%; Pred. No. 6e-28;
Matches 184; Conservative 46; Mismatches 102; Indels 95; Gaps 19;
QY 14 LEHRFGLDYSSLY-MCKPRSMKR-DDTKVSDTYKLPRLIEKKRRDRINECIAQLKD 71
DB 20 LEHGLSGMDFAHMYQVYKSRGIRKSRDSK--ETKLPRLIEKKRRDRINECIAQLKD 77
QY 72 LLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQOQHOKIIALQNGERS---LKSPQIS 128
DB 78 LLPEHLKLTTLGHLEKAVVLETLKHLKALTALTEQOQHOKIIALQNGLSGRNLEA 137
QY 129 DLDAPHSGFQCAKEVLYLSRFESWTPREPCVOLINHLHAVATQFLPTQLTQQVPL 188
DB 138 GQEMFCSGFQTCAREVLYLAKHEN--TRDLKSSQLVTHLRVVSSELL---QGCASRKFL 192
QY 189 SKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPIQRT-OPSAEALAAENDTDTDSYGGE 247
DB 193 D---SAPKAVDLKPKPLANGSE-GPGKNCVPIQRTFAPSGGEGSGSDTDTDSYGGE 248
QY 248 AEA-----RPDREKKGAGASRVTIKQEPPEGSDSPAPK-RMKLDGRGGSGGPG 296
DB 249 LEKGLRSEQPYFKSDHGRRAVGERVSTIKQE--SEPTKKSRMQLSEEGHFAG--- 303
QY 297 GGAATAAALLGPPDPAALRRPDAALLSSLVAFGGGGAPPQPAAAAAPPCFLP 356
DB 304 -----SDLM-----GSPFLGPHPHQPPFCLPF-YL 327
QY 357 SPSAAAAYVOPFLDKSGLEKLYLPAAAAAPPFLYPCGPAPAAAAA-----AFAFC 416
DB 328 IPPSATAYL-PM-----LEKCVPTSV-----PVLVPLGNTSAAA----- 361
QY 417 LSSVLP 423
DB 362 LSSFPNP 368
RESULT 10
BAC41056
ID BAC41056 PRELIMINARY; PRT; 411 AA.
AC BAC41056;
DT 14-APR-2004 (TrEMBLrel. 27, Created)
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)
DE 14-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Male submandibular gland CRU-1734 SCA-9 clone 15 cDNA, RIKEN
DE full-length enriched library, clone:G430067G17 product:basic
DE helix-loop-helix domain containing, class B2, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Submandibular gland;
RX MEDLINE=22354683; PubMed=12466851;
RA The RIKEN Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Submandibular gland;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 420:563-573 (2002).

```

RL	Nature 409:685-690 (2001).	297	CGAAAAAALLGPDPAAPAAAALLRPAALLSSIVAFGGGGGAPFPAPAAAAAAPPCLPFCFL	356
RN	[3]	304	-----SDLM-----GSPFLGPHQPPFCLPF-YL	327
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Swiss Webster; TISSUE=Submandibular gland;			
RX	MEDLINE=99279253; PubMed=10349636;			
RA	Carninci P., Hayashizaki Y.;			
RT	"High-efficiency full-length cDNA cloning.";			
RL	Meth. Enzymol. 303:19-44 (1999).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Swiss Webster; TISSUE=Submandibular gland;			
RX	MEDLINE=20499374; PubMed=11042159;			
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,			
RT	Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;			
RL	"Normalization and subtraction of cap-trapper-selected cDNAs to			
RP	prepare full-length cDNA libraries for rapid discovery of new genes.";			
RN	Genome Res. 10:1617-1630 (2000).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Swiss Webster; TISSUE=Submandibular gland;			
RX	MEDLINE=20530913; PubMed=11076861;			
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,			
RT	Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,			
RL	Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,			
RP	Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,			
RT	Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,			
RL	Okazaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,			
RP	Yoneda Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;			
RT	"RIKEN integrated sequence analysis (RISA) system-384-format			
RL	sequencing pipeline with 384 multicapillary sequencer.";			
RN	Genome Res. 10:1757-1771 (2000).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Swiss Webster; TISSUE=Submandibular gland;			
RA	Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,			
RT	Pukada S., Furuno M., Hanagaki T., Hara A., Hashizume W.,			
RL	Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,			
RP	Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,			
RT	Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,			
RL	Nishihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,			
RP	Rurihi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,			
RT	Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,			
RL	Tagawa A., Takahashi F., Takaku-Akahira S., Takada Y., Tagami M.,			
RP	Tomaru A., Toya T., Yasuishi A., Muramatsu M., Hayashizaki Y.;			
RT	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AK090032; BAC41056.1; -			
SQ	SEQUENCE 411 AA; 45360 MW; B392893CD49292BC CRC64;			
	Query Match 27.2%; Score 685.5; DB 2; Length 411;			
	Best Local Similarity 43.1%; Pred. No. 6e-28;			
	Matches 184; Conservative 46; Mismatches 102; Indels 95; Gaps 19;			
QY	14 LEHDFGLDYSSLY-MCKPKRMKR-DDTKVSDTYKLPHRLIEKRRDRINECTAQLKD 71			
DB	20 LEHGLDGLMDFAHMYQVYKSRGIRKSRSDSK--ETVYKLPRLIEKRRDRINECTAQLKD 77			
QY	72 LLPEHLKLTTLGHLKAVLELTLKHLKALTALTEQOHOKIALONGERS---LKSPIQS 128			
DB	78 LLPEHLKLTTLGHLKAVLELTLKHLKALTALTEQOHOKIALONGERS---LKSPIQS 137			
QY	129 DLDAFHSGFGTCAKEVLQYLRSFESWTPREPCVQLINHLHAYATVFLPTPQLLTQVPL 188			
DB	138 GQEMFCGSGFTCAREVLYLAKHEN--TRDLKSSQLVTLHLRVVSELL---QGGASRKPL 192			
QY	189 SKGTGAPSAGSAAAPCLERAGQKLEPLAYCVPIQRT-OPSALAEADTTDSCGYGE 247			
DB	193 D---SAPKAVDLKPKSFLAKGSE-GEKKNVCVPIQRTFAPSQGGSGSDTTDSCGYGE 248			
QY	248 AEA-----RPDKKGGKAGASRTTIKQPPGSDSPAPK-RMKLDSRGCGSGGPG 296			
DB	249 LEKGDLSRQEPYFKSDHGRFPAVGERVSTIKQE--SEEPPTKSRMQLSEEGHFAG--- 303			
	Query Match 27.2%; Score 684.5; DB 2; Length 403;			
	Best Local Similarity 40.1%; Pred. No. 6.6e-28;			
	Matches 186; Conservative 53; Mismatches 122; Indels 103; Gaps 19;			
QY	6 PHLQERQLLEHDFGLDYSSLYMCKPKRMKR-DDTKVSDTYKLPHRLIEKRRDRINE 64			

Db 11 PCMKHPSLSDMQGDF-PMVYKPRGKMSK--DTYKLPRLIEKKRRDRINE 67
QY 65 CIAQLKDLLPEHLKLTTLGLHLEKAVLELTTLKHLKALTALTEQHQKIIALQN-----GER 120
Db 68 CIAQLKDLLPEHLKLTTLGLHLEKAVLELTTLKHLKALTALTEQHQKIIALQN-----GER 127
QY 121 SLKSPQSDLDLAFHSQFQTCQKAVLQVLSFESWTREPVCVQLINHLHVAVATQFLPTQ 180
Db 128 G-NGPSENSEMFSGHLCQKAVLQFLANQE--TWRLDTTAHIIIEHLQKVASLQSP- 183
QY 181 LLTQQVPLSKGTGAPSAAGSAAPCLERAGOK---LEPLA-----YCVPIQRTOPSAE 231
Db 184 -----PSRLDEPAKQESREKPSGLQPKAEGHAKNCVPIQRTYPHSS 229
QY 232 LAENDTDTSGYGGGAERAPDRE-----KKGAGASRVTIKQPPGEDSPAPKMKL 284
Db 230 EQSGSDTDTSGYGGYERDQKQAPDCYVKESGALKYSSIKEE---QDEPPSKRPRS 286
QY 285 DSRGGSGGGPGGGAAGAAALGPDPAALLLRPDAALLSLVAFGGGGAPF----- 339
Db 287 DS-----SEDESLSGHVV-----GHSYVVSFSP 311
QY 340 PQPAAAAAFCPLPFLSPSAAAAYVQFLDKSGLEKLYLPAAAAFPFLPYGIPAPAA 399
Db 312 PQ-----PLCMPFYLPPGAAAYL-PM-----LEKWPYGA---MPVLYPGLSSPA 355
QY 400 AAAAAAATAAFAAFCPLSSVSPPEKAGAAATLLPHEVAPL 443
Db 356 SLSEKPLPSSWMSSRVGSPVSTPTSDMSFALLQAL--KQVPPL 397

RESULT 12
AAH66738
ID AAH66738 PRELIMINARY; PRT; 403 AA.
AC AAH66738
DT 01-JUN-2004 (TEMBLrel. 27, Created)
DT 01-JUN-2004 (TEMBLrel. 27, Last sequence update)
DE Hypothetical protein zgc:76913.
GN ZGC:76913.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., S.W.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066738; AAH66738.1; --
KW Hypothetical protein.
SQ SEQUENCE 403 AA; 44682 MW; 08442DC03FF298E CRC64;

Query Match 27.2%; Score 684.5; DB 2; Length 403;
Best Local Similarity 40.1%; Pred. No. 6.6e-28;
Matches 186; Conservative 53; Mismatches 122; Indels 103; Gaps 19;

QY 6 PHLOERQLLEHRDFGLDYSLLYNCKPKSKMKR-DDTKVSDTYKLPRLIEKKRRDRINE 64
Db 11 PCMKHPSLSDMQGDF-PMVYKPRGKMSK--DTYKLPRLIEKKRRDRINE 67
QY 65 CIAQLKDLLPEHLKLTTLGLHLEKAVLELTTLKHLKALTALTEQHQKIIALQN-----GER 120
Db 68 CIAQLKDLLPEHLKLTTLGLHLEKAVLELTTLKHLKALTALTEQHQKIIALQN-----GER 127
QY 121 SLKSPQSDLDLAFHSQFQTCQKAVLQVLSFESWTREPVCVQLINHLHVAVATQFLPTQ 180
Db 128 G-NGPSENSEMFSGHLCQKAVLQFLANQE--TWRLDTTAHIIIEHLQKVASLQSP- 183
QY 181 LLTQQVPLSKGTGAPSAAGSAAPCLERAGOK---LEPLA-----YCVPIQRTOPSAE 231
Db 184 -----PSRLDEPAKQESREKPSGLQPKAEGHAKNCVPIQRTYPHSS 229
QY 232 LAENDTDTSGYGGGAERAPDRE-----KKGAGASRVTIKQPPGEDSPAPKMKL 284
Db 230 EQSGSDTDTSGYGGYERDQKQAPDCYVKESGALKYSSIKEE---QDEPPSKRPRS 286
QY 285 DSRGGSGGGPGGGAAGAAALGPDPAALLLRPDAALLSLVAFGGGGAPF----- 339
Db 287 DS-----SEDESLSGHVV-----GHSYVVSFSP 311
QY 340 PQPAAAAAFCPLPFLSPSAAAAYVQFLDKSGLEKLYLPAAAAFPFLPYGIPAPAA 399
Db 312 PQ-----PLCMPFYLPPGAAAYL-PM-----LEKWPYGA---MPVLYPGLSSPA 355
QY 400 AAAAAAATAAFAAFCPLSSVSPPEKAGAAATLLPHEVAPL 443
Db 356 SLSEKPLPSSWMSSRVGSPVSTPTSDMSFALLQAL--KQVPPL 397

RESULT 13
BHB2 RAT
ID BHB2 RAT STANDARD; PRT; 411 AA.
AC O35780;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Class B basic helix-loop-helix protein 2 (bHLH2) (Enhancer-of-split
and hairy-related protein 2) (SHARP-2).
GN Name=bhlh2; Synonyms=Sharp2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
RX MEDLINE=98193761; PubMed=9532582;
RA Rossier M.J., Doerr J., Gass P., Schwab M.H., Nave K.-A.;
RT "SHARPs: mammalian enhancer-of-split- and hairy-related proteins
coupled to neuronal stimulation";
RL Mol. Cell. Neurosci. 10:460-475 (1997).
CC - FUNCTION: May function as a transcriptional factor for neuronal
differentiation.
CC - SUBUNIT: Able to homodimerize or heterodimerize with E47 (By
similarity).
CC - SUBCELLULAR LOCATION: Nuclear (By similarity).
CC - TISSUE SPECIFICITY: Expressed in heart, brain, spleen, lung,
CC liver, muscle, kidney, uterus and gut.
CC - SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC - SIMILARITY: Contains 1 orange domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF009330; RAB63587.1; --
DR RGD: 68439; Bhlhb2.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003650; Orange.
DR Pfam: PF07527; HairY_Orange; 1.
DR Pfam: PF000107; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00511; ORANGE; 1.
DR PROSITE: PS00888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DNA BIND 53 65 Basic motif.
FT DOMAIN 66 108 Helix-loop-helix motif.
FT DOMAIN 140 184 Orange.
SQ SEQUENCE 411 AA; 45529 MW; E56BD468D08824AD CRC64;

Query Match 27.1%; Score 682.5; DB 1; Length 411;
Best Local Similarity 43.5%; Pred. No. 8.6e-28;
Matches 186; Conservative 45; Mismatches 100; Indels 97; Gaps 20;

QY 14 LEHRDFIGLDYSSLY-MCKPKRSMKR-DTKVSDTYKLPHRLIEKRRDRINECIAQLKD 71
DB 20 LEHGDLSGMDFAHMYQVYKSRGKESDSK--ETVKLPHRLIEKRRDRINECIAQLKD 77
QY 72 LPEHLKLTTLGLHLEKAVVLELTALKALTALEQHOHOKIALQN-----GERSLKSPIQ 127
DB 78 LPEHLKLTTLGLHLEKAVVLELTALKALTALEQHOHOKIALQN-----GERSLKSPIQ 136
QY 128 SDLDAPHSGFOTCAKEVLYLSRFSWTPREPRCVQLINHLHAVATQFLTPQLLTQQVVP 187
DB 137 AGEMFCGSGFOTCAREVLYLAKHEN--TRDLKSSQLVTHLRVVSSELL---QGSASRKP 191
QY 188 LSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPIQRT-OPSAEIAAENDTDTDSGYGG 246
DB 192 LD---SAPKEVDPEKPSFLAKGSE-GPGKNCVPIQRTFAPSGGSGSDTDTDSGYGG 247
QY 247 EABA-----RPDREKKGAGASRVTIKQEPGDSBPAPK-RMKLDSRGSGSGGSP 295
DB 248 ELEKGLDLSQPYFKSDHGRFTVGERVSTIKQE--SEEPPTKSRMQLSDEEGHFGV-- 303
QY 296 GGGAAAAAALLGPDPAALALRPDAALLSSLVAFGGGGAPFPQAAAAAPFCLPFCF 355
DB 304 -----SDLM-----GSPFLGPHPHQPPFCLPF-Y 326
QY 356 LSPSAAAAYVQFLDKSGLEKLYPAAAAAPFLLYGPAPAAAAAALAAAAAAPP 415
DB 327 LIPPSATAYL-PM-----LEKWCYPTSV-----PLLYPGLNTSAAA----- 361
QY 416 CLSVLSP 423
DB 362 -LSSFWNP 368

RESULT 14
QY Q76JQ4 PRELIMINARY; PRT; 411 AA.
AC Q76JQ4
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split- and hairy-related protein-2.
GN Name=SHARP-2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=14725860;
RA Hirano S., Yamada K., Kawata H., Shou Z., Mizutani T., Shigematsu Y.,
RA Mayumi M., Miyamoto K.,
RT "The rat enhancer of split- and hairy-related protein-2 gene: hepatic
RT expression, genomic structure, and promoter analysis.",
RL Arch. Biochem. Biophys. 422:81-90(2004).
DR EMBL: AB096137; BAD01588.1; --
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003650; Orange.
DR Pfam: PF07527; HairY_Orange; 1.
DR Pfam: PF000107; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00511; ORANGE; 1.
DR PROSITE: PS00888; HLH; 1.
SQ SEQUENCE 411 AA; 45557 MW; 2086DABB3DB82058 CRC64;

Query Match 27.0%; Score 679.5; DB 2; Length 411;
Best Local Similarity 43.1%; Pred. No. 1.2e-27;
Matches 184; Conservative 45; Mismatches 103; Indels 95; Gaps 19;

QY 14 LEHRDFIGLDYSSLY-MCKPKRSMKR-DTKVSDTYKLPHRLIEKRRDRINECIAQLKD 71
DB 20 LEHGDLSGMDFAHMYQVYKSRGKESDSK--ETVKLPHRLIEKRRDRINECIAQLKD 77
QY 72 LPEHLKLTTLGLHLEKAVVLELTALKALTALEQHOHOKIALONGERS---LKSPIOS 128
DB 78 LPEHLKLTTLGLHLEKAVVLELTALKALTALEQHOHOKIALONGERS---LKSPIOS 137
QY 129 DLDAFHSGFOTCAKEVLYLSRFSWTPREPRCVQLINHLHAVATQFLTPQLLTQQVPL 188
DB 138 GQEMFCGSGFOTCAREVLYLAKHEN--TRDLKSSQLVTHLRVVSSELL---QGSASRKP 192
QY 189 SKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPIQRT-OPSAEIAAENDTDTDSGYGG 247
DB 193 D---SAPKEVDPEKPSFLAKGSE-GPGKNCVPIQRTFAPSGGSGSDTDTDSGYGG 248
QY 248 EABA-----RPDREKKGAGASRVTIKQEPGDSBPAPK-RMKLDSRGSGSGGSP 296
DB 249 LEKGLDLSQPYFKSDHGRFTVGERVSTIKQE--SEEPPTKSRMQLSDEEGHFGV-- 303
QY 297 GGGAAAAAALLGPDPAALALRPDAALLSSLVAFGGGGAPFPQAAAAAPFCLPFCFL 356
DB 304 -----SDLM-----GSPFLGPHPHQPPFCLPF-YL 327
QY 357 LSPSAAAAYVQFLDKSGLEKLYPAAAAAPFLLYGPAPAAAAAALAAAAAAPP 416
DB 328 LIPPSATAYL-PM-----LEKWCYPTSV-----PLLYPGLNTSAAA----- 361
QY 417 LSVLSP 423
DB 362 LSSFWNP 368

RESULT 15
QY Q76JQ4 PRELIMINARY; PRT; 411 AA.
AC Q76JQ4
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Enhancer of split- and hairy-related protein-2.
GN SHARP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: December 15, 2004, 20:05:13 ; Search time 40 Seconds
(without alignments)
1164.223 Million cell updates/sec

Title: US-10-078-650-12
Perfect score: 484
Sequence: 1 MDSGIHLQERQLLEHRDFI.....PGNPSSAQEDSPQPKKEAP 484

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	440	90.9	482	JC7583	basic helix-loop-h
2	81	16.7	410	JC7584	basic helix-loop-h
3	52	10.7	412	JC5547	basic helix-loop-h
4	16	3.3	305	I57039	genomic screen hom
5	16	3.3	314	JC5273	paired type homeob
6	16	3.3	403	A53662	homeotic protein H
7	16	3.3	495	S31223	transcription fact
8	16	3.3	513	A48233	polyomavirus enhan
9	16	3.3	627	T02610	probable YME1 ATP-
10	16	3.3	1028	A56038	DNA-binding protei
11	16	3.3	1213	S16356	ovo protein - frui
12	16	3.3	1668	T13748	sex comb protein -
13	15	3.1	220	JC5954	ribosomal protein
14	15	3.1	401	A48423	engrailed homeodm
15	15	3.1	420	T14911	bZIP DNA-binding p
16	15	3.1	443	I38239	transcription fact
17	15	3.1	796	JC7555	Clorf4 protein -
18	15	3.1	806	T13690	hypothetical prote
19	15	3.1	1065	T13230	dachshund isoform
20	15	3.1	1072	T13232	dachshund protein
21	15	3.1	1074	T13229	dachshund protein
22	15	3.1	1081	T13231	dachshund protein
23	15	3.1	1205	A55015	bumetanide-sensiti
24	15	3.1	1212	A57187	bumetanide-sensiti
25	15	3.1	2639	T31328	fibroin - Chinese
26	14	2.9	150	T12547	hypothetical prote
27	14	2.9	289	A43562	homeotic protein H
28	14	2.9	334	G02409	protein kinase C-b
29	14	2.9	364	I48188	Gene NKx6.1 protei

30 14 2.9 374 2 T03875 probable homeobox
31 14 2.9 375 2 T03874 probable homeobox
32 14 2.9 494 2 A42170 zinc finger protei
33 14 2.9 552 1 WJFFEN homeotic protein e
34 14 2.9 828 2 C98402 protein HUSC05.1
35 14 2.9 1355 2 S40022 spalt protein - fr
36 14 2.9 1506 2 T51900 related to PAN2 pr
37 14 2.9 2038 2 A43742 female sterile hom
38 13 2.7 233 2 S11563 probable MASH-2 pr
39 13 2.7 238 2 A48279 achaete acute prot
40 13 2.7 323 2 S16318 homeotic protein H
41 13 2.7 331 2 B47236 zinc-finger protei
42 13 2.7 376 2 A26086 segmentation prote
43 13 2.7 378 2 A44433 basic helix-loop-h
44 13 2.7 477 2 A47236 zinc-finger protei
45 13 2.7 497 2 JC5076 myc-associated zin

ALIGNMENTS

RESULT 1

JC7583
basic helix-loop-helix protein, DEC2 - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7583
Rifujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.
Biochem. Biophys. Res. Commun. 280, 164-171, 2001
A:Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop
A:Reference number: JC7583; MUID:21092582; PMID:11162494
A:Accession: JC7583
A:Molecule type: mRNA
A:Residues: 1-482 <FU>
A:Cross-references: UNIPROT:Q9C0J9; DDBJ:AB044088
C:Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix p
C:Genetics: C:Genetics:
A:Map position: 12p11.23-p12.1
C:Keywords: transcription factor
F:1-173/Region: highly conserved #status predicted
F:130-173/Domain: Orange #status predicted <ORA>
F:286-411/Region: alanine and glycine-rich #status predicted

Query Match	90.9%	Score	440	DB	2	Length	482
Best Local Similarity	100.0%	Pred. No.	0				
Matches	440	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	45	DTYKLPRLIEKKERDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTIKHLKALTAL	104				
Db	43	DTYKLPRLIEKKERDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTIKHLKALTAL	102				
QY	105	TEQHQKIIALONGERSLKSPISQDLDAFHSGFQTCAKEVLOYLSRESWTPEPRCVOL	164				
Db	103	TEQHQKIIALONGERSLKSPISQDLDAFHSGFQTCAKEVLOYLSRESWTPEPRCVOL	162				
QY	165	INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPIQ	224				
Db	163	INHLHAVATQFLPTPQLLTQQVPLSKGTGAPSAAGSAAAPCLERAGOKLEPLAYCVPIQ	222				
QY	225	RTQPSAELAENDTDTDSGGGGAEEARPDPREKKGAGASRVTIKQEPFGGSDSPAPKMKL	284				
Db	223	RTQPSAELAENDTDTDSGGGGAEEARPDPREKKGAGASRVTIKQEPFGGSDSPAPKMKL	282				
QY	285	DSRGGGGGGGGGAAAAAALLGPDPAALLLRPDAALLSLVAFGGGGGAPFPPOPAA	344				
Db	283	DSRGGGGGGGGGAAAAAALLGPDPAALLLRPDAALLSLVAFGGGGGAPFPPOPAA	342				
QY	345	AAAPFCPLPFCFLSPSAAAAAYVQPFLLDKSGLEKLYPAAAAAPFPPLLYPGIPAPAAAAAAA	404				
Db	343	AAAPFCPLPFCFLSPSAAAAAYVQPFLLDKSGLEKLYPAAAAAPFPPLLYPGIPAPAAAAAAA	402				
QY	405	AAAAAAAAPFCLSSVLSPPPEKAGAAAATLLPHEVAPLGPAPHPQHPGHTLPPFAGPRE	464				

Db 403 AAAAAAAAAA... 462

QY 465 PGNPSSAEDPSQPKKEAP 484

Db 463 PGNPSSAEDPSQPKKEAP 482

RESULT 2

JC7584

basic helix-loop-helix protein, DEC2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004

C:Accession: JC7584

R:Fujimoto, K.; Shen, M.; Noshiro, M.; Matsubara, K.; Shingu, S.; Honda, K.; Yoshida, E.

Biochem. Biophys. Res. Commun. 280, 164-171, 2001

A:Title: Molecular cloning and characterization of DEC2, a new member of basic helix-loop-helix protein family

A:Reference number: JC7583; MUID:21092582; PMID:11162494

A:Accession: JC7584

A:Molecule type: mRNA

A:Residues: 1-410 <FUJ>

A:Cross-references: UNIPROT:Q99PV5; DBJ:AB044090

C:Comment: This protein, a novel member of the DEC subfamily of basic helix-loop-helix proteins

C:Genetics:

A:Gene: dec2

A:Map position: 6 G2-G3

C:Keywords: transcription factor

F:1-173/Region: highly conserved #status predicted

F:130-173/Domain: Orange #status predicted <ORA>

F:277-344/Region: alanine and glycine-rich #status predicted

Query Match 16.7%; Score 81; DB 2; Length 410;

Best Local Similarity 100.0%; Pred. No. 2.9e-64;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 DTVKLPHRLIEKKRRDRINICIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHALTAL 104

Db 43 DTVKLPHRLIEKKRRDRINICIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKHALTAL 102

QY 105 TEQHQKIIALONGERSLKSP 125

Db 103 TEQHQKIIALONGERSLKSP 123

RESULT 3

JC5547

basic helix-loop-helix factor DEC1 - human

C:Species: Homo sapiens (man)

C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004

C:Accession: JC5547

R:Shen, M.; Kawamoto, T.; Yan, W.; Nakanasu, K.; Tamagami, M.; Koyano, Y.; Noshiro, M.;

Biochem. Biophys. Res. Commun. 236, 294-298, 1997

A:Title: Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in human embryonic kidney cells

A:Reference number: JC5547; MUID:97382424; PMID:9240428

A:Accession: JC5547

A:Molecule type: mRNA

A:Residues: 1-412 <SHE>

A:Cross-references: UNIPROT:O14503; DBJ:AB004066; NID:G2308996; PIDN:BAA21720.1; PID:G2308996

C:Comment: This protein is involved in the control of cell differentiation in several tissues

F:51-108/Region: helix-loop-helix #status predicted

Query Match 10.7%; Score 52; DB 2; Length 412;

Best Local Similarity 100.0%; Pred. No. 2.4e-38;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 TYKLPRLIEKKRRDRINICIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKH 97

Db 52 TYKLPRLIEKKRRDRINICIAQLKDLLPEHLKLTTLGHLEKAVVLELTTLKH 103

RESULT 4

J57039

genomic screen homeobox protein 2 - mouse

C:Species: Mus sp. (mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Aug-2004

C:Accession: J57039

R:Hsieh-Li, H.M.; Witte, D.P.; Szucsik, J.C.; Weinstein, M.; Li, H.; Potter, S.S.

Mech. Dev. 50, 177-186, 1995

A:Title: Gsh-2, a murine homeobox gene expressed in the developing brain.

A:Reference number: J57039; MUID:95344993; PMID:7619729

A:Accession: J57039

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-305 <RES>

A:Cross-references: GB:S79041; NID:G1042008; PIDN:AAB34947.1; PID:G1042009

C:Genetics:

A:Gene: Gsh-2

C:Superfamily: homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:204-260/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 2; Length 305;

Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAAAAAAAAA 413

Db 147 AAAAAAAAAAAAAAAAAAAAAA 162

RESULT 5

JC5273

paired type homeobox protein, NBP - human

C:Species: Homo sapiens (man)

C:Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 16-Aug-2004

C:Accession: JC5273

R:Fukuyama, M.; Nishi, Y.; Yoshii, J.; Okubo, K.; Matsubara, K.

DNA Res. 3, 311-320, 1996

A:Title: Identification and cloning of neuroblastoma-specific and nerve tissue-specific homeobox genes

A:Reference number: JC5272; MUID:97191543; PMID:9039501

A:Contents: neuroblastoma cell

A:Accession: JC5273

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-314 <YOK>

A:Cross-references: UNIPROT:Q99453; DBJ:D82344; NID:G1841337; PIDN:BAA1555.1; PID:G1841337

C:Comment: This protein is a transcriptional repressor involved in regulating gene expression

C:Superfamily: homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:99-155/Domain: homeobox homology <HOX>

Query Match 3.3%; Score 16; DB 2; Length 314;

Best Local Similarity 100.0%; Pred. No. 2.9e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAAAAAAAAA 413

Db 241 AAAAAAAAAAAAAAAAAAAAAA 256

RESULT 6

AS3662

homeotic protein HB9 - human

C:Species: Homo sapiens (man)

C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 17-Oct-1997

C:Accession: AS3662

R:Harrison, K.A.; Druey, K.M.; Deguchi, Y.; Tusciano, J.M.; Kehrl, J.H.

J. Biol. Chem. 269, 19968-19975, 1994

A:Title: A novel human homeobox gene distantly related to proboscipedia is expressed in the developing brain

A:Reference number: AS3662; MUID:94327547; PMID:7914194

A:Accession: AS3662

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-403 <HAR>

A:Cross-references: GB:U07663

A>Note: the nucleotide sequence and conceptual translation as given are self-consistent

C;Genetics:
A;Gene: GDB:HLXB9
A;Cross-references: GDB:136411; OMIM:142994
A;Map position: 1q41-1q42.1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;244-300/Domain: homeobox homology <Hox>

Query Match 3.3%; Score 16; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
DB 121 AAAAAAAAAAAAAA 136

RESULT 7
S31223
transcription factor Brn-1 - mouse
N;Alternate names: class III POU domain protein brain-1
C;Species: Mus musculus (house mouse)
C;Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S31223
R;Hara, Y.; Roverscailli, A.C.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 3280-3284, 1992
A;Title: Structure and evolution of four POU domain genes expressed in mouse brain.
A;Reference number: S31223; MUID:92228768; PMID:1565620
A;Accession: S31223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-495 <HAR>
A;Cross-references: UNIPROT:P31361; EMBL:M88295; NID:G200444; PIDN:AAA39960.1; PID:G20044
C;Superfamily: transcription factor Brn-1; homeobox homology; POU domain homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;101-112/Region: alanine-rich
F;162-180/Region: histidine/proline-rich
F;186-201/Region: alanine-rich
F;236-247/Region: glycine-rich
F;267-291/Region: histidine/proline-rich
F;316-383/Domain: POU domain homology <POU>
F;402-458/Domain: homeobox homology <Hox>

Query Match 3.3%; Score 16; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
DB 186 AAAAAAAAAAAAAA 201

RESULT 8
A48233
polyomavirus enhancer-binding protein 2 alpha chain type 1 - mouse
N;Alternate names: PEA2 alpha chain type 1; PEA2 alpha chain type 2; PEBP2 alpha chain t
C;Species: Mus musculus (house mouse)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A48233; B48233
R;Ogawa, E.; Maruyama, M.; Kagoshima, H.; Inuzuka, M.; Lu, J.; Satake, M.; Shigesada, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 6859-6863, 1993
A;Title: PEBP2/PEA2 represents a family of transcription factors homologous to the produ
A;Reference number: A48233; MUID:93342088; PMID:8341710
A;Accession: A48233
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-513 <OGA>
A;Cross-references: UNIPROT:Q08775; GB:D14636; NID:G931766; PIDN:BAA03485.1; PID:G100399
A;Accession: B48233
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-304, 'L', 306 <OG2>

A;Cross-references: GB:D14637; NID:G391768; PIDN:BAA03486.1; PID:G391769
C;Genetics:
A;Gene: PEBP2alpha
A;Superfamily: transcription factor CBF alpha 2
C;Keywords: alternative splicing; DNA binding; T-cell; transcription factor; transcripti
F;244-300/Domain: homeobox homology <Hox>

Query Match 3.3%; Score 16; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
DB 64 AAAAAAAAAAAAAA 79

RESULT 9
T02610
probable YME1 ATP-dependent proteinase [imported] - Arabidopsis thaliana
N;Alternate names: YTA1 protein homolog T19L18.5
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02610; H84656
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, August 1998
A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.
A;Reference number: Z14681
A;Accession: T02610
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-627 <ROU>
A;Cross-references: UNIPROT:O80983; EMBL:AC004747; NID:G3413696; PID:G3413700
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84656
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-627 <STO>
A;Cross-references: GB:AE002093; NID:G3413700; PIDN:AAC31223.1; GSPDB:GN00139
C;Genetics:
A;Gene: T19L18.5; At2g26140
A;Map position: 2
A;Introns: 40/1; 119/1; 160/3; 319/3
C;Superfamily: cell division protein ftsh; Ftsh/SEC18/CDC48-type ATP-binding domain homo

Query Match 3.3%; Score 16; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAA 413
DB 598 AAAAAAAAAAAAAA 613

RESULT 10
A56038
DNA-binding protein ovo - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: A56038
R;Garfinkel, M.D.; Wang, J.; Liang, Y.; Mahowald, A.P.
Mol. Cell. Biol. 14, 6809-6818, 1994
A;Title: Multiple products from the shavenbaby-ovo gene region of Drosophila melanogaste
A;Reference number: A56038; MUID:95021209; PMID:7935398
A;Accession: A56038
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1028 <GAR>
A;Cross-references: UNIPROT:P51521; GB:U11383; NID:G520526; PIDN:AAB60216.1; PID:G520527

```
C;Genetics:
A;Gene: ovo
A;Cross-references: FlyBase:FBgn0003028

Query Match      3.3%; Score 16; DB 2; Length 1028;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAA 413
Db 497 AAAAAAAAAAAAAAA 512

RESULT 11
S16356
ovo protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S16356
R;Mevel-Ninio, M.; Terracol, R.; Kafatos, F.C.
EMBO J. 10, 2259-2266, 1991
A;Title: The ovo gene of Drosophila encodes a zinc finger protein required for female sex
A;Reference number: S16356; MUID:91293102; PMID:1712294
A;Accession: S16356
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1213 <MEV>
A;Cross-references: UNIPROT:Q8T8L9; EMBL:X59772
C;Genetics:
A;Gene: FlyBase:ovo
A;Cross-references: FlyBase:FBgn0003028
A;Introns: 931/3; 1152/3

Query Match      3.3%; Score 16; DB 2; Length 1213;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAA 413
Db 860 AAAAAAAAAAAAAAA 875

RESULT 12
T13748
sex comb protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13748
R;Sinclair, D.A.R.; Milne, T.A.; Hodgson, J.W.; Shellard, J.; Salinas, C.A.; Kyba, M.; R
Development 125, 1207-1216, 1998
A;Title: The additional sex combs gene of Drosophila encodes a chromatin protein that bi
A;Reference number: Z17750; MUID:98146384; PMID:9477319
A;Accession: T13748
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1668 <GIN>
A;Cross-references: EMBL:AJ001164; NID:G3292938; PIDN:CAA04568.1; PID:G3292939
C;Genetics:
A;Cross-references: FlyBase:FBgn000142
C;Function:
A;Description: involved in repression of homeotic loci

Query Match      3.3%; Score 16; DB 2; Length 1668;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAA 413
Db 129 AAAAAAAAAAAAAAA 144

RESULT 13
JC5954
ribosomal protein L14 - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
A;Accession: JC5954
R;Tanaka, M.; Tanaka, T.; Harata, M.; Suzuki, T.; Mitsu, Y.
Biochem. Biophys. Res. Commun. 243, 531-537, 1998
A;Title: Triplet repeat-containing ribosomal protein L14 gene in immortalized human endot
A;Reference number: JC5954; MUID:98153799; PMID:9480843
A;Accession: JC5954
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <TAN>
A;Cross-references: DDBJ:D87735; NID:G1620021; PIDN:BAA13443.1; PID:G1620022
C;Superfamily: rat ribosomal protein L14

Query Match      3.1%; Score 15; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 AAAAAAAAAAAAAAA 412
Db 150 AAAAAAAAAAAAAAA 164

RESULT 14
A48423
engrailed homeodomain-containing protein En-1 - mouse
N;Alternate names: homeotic protein En-1
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A48423; S13009; A26629; A24778
R;Logan, C.; Hanks, M.C.; Noble-Topham, S.; Nallainathan, D.; Provart, N.J.; Joyner, A.L.
Dev. Genet. 13, 345-358, 1992
A;Title: Cloning and sequence comparison of the mouse, human, and chicken engrailed genes
A;Reference number: A48423; MUID:93185339; PMID:1363401
A;Accession: A48423
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-401 <LOG>
A;Cross-references: UNIPROT:P09065
A;Experimental source: CD-1, embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:126620)
R;Holland, P.W.H.; Williams, N.A.
FEBS Lett. 277, 250-252, 1990
A;Title: Conservation of engrailed-like homeobox sequences during vertebrate evolution.
A;Reference number: S13009; MUID:9109509; PMID:1980115
A;Accession: S13009
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 321-380 <HOL>
R;Joyner, A.L.; Martin, G.R.
Genes Dev. 1, 29-38, 1987
A;Title: En-1 and En-2, two mouse genes with sequence homology to the Drosophila engrail
A;Reference number: A91620; MUID:88112776; PMID:2892757
A;Accession: A26629
A;Molecule type: DNA; mRNA
A;Residues: 278-401 <JOY>
A;Cross-references: GB:Y00201; GB:M11987; NID:G49587; PIDN:CAA68361.1; PID:G669105
R;Joyner, A.L.; Kornberg, T.; Coleman, K.G.; Cox, D.R.; Martin, G.R.
Cell 43, 29-37, 1985
A;Title: Expression during embryogenesis of a mouse gene with sequence homology to the D
A;Reference number: A24778; MUID:86079501; PMID:2416459
A;Accession: A24778
A;Molecule type: DNA
A;Residues: 311-401 <JO2>
C;Genetics:
A;Gene: en-1
A;Map position: 1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;313-369/Domain: homeobox homology <HOX>

Query Match      3.1%; Score 15; DB 2; Length 401;
```

Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 PAAAAAAAAAAAAA 411
 |||||
 Db 206 PAAAAAAAAAAAAA 220

RESULT 15
 T14911
 bZIP DNA-binding protein - parsley
 C:Species: Petroselinum crispum (parsley)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T14911
 R:Kircher, S.; Ledger, S.; Hayashi, H.; Weishaar, B.; Schafer, E.; Frohnmeyer, H.
 Mol. Gen. Genet. 257, 595-605, 1998
 A>Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis of
 A:Reference number: Z18261; MUID:98265918; PMID:9604882
 A:Accession: T14911
 A>Status: preliminary; translated from GE/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-420 <KIR>
 A:Cross-references: UNIPROT:O82037; EMBL:Y10810; NID:g3336904; PIDN:CAA71770.1; PID:g333
 C:Genetics:
 A:Note: CPRF4b
 C:Superfamily: fava bean G-box-binding protein; fos/jun DNA-binding domain homology
 C:Keywords: DNA binding; leucine zipper

Query Match 3.1%; Score 15; DB 2; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 397 PAAAAAAAAAAAAA 411
 |||||
 Db 30 PAAAAAAAAAAAAA 44

Search completed: December 15, 2004, 20:15:07
 Job time : 41 secs

This Page Blank (uspto)